Reputation and Reciprocity

Dissertation
for the Faculty of Economics, Business Administration and Information Technology of the University of Zürich

to achieve the title of
Doctor of Economics

presented by

Frédéric-Guillaume Schneider

from Germany and France

approved at the request of

Prof. Dr. Ernst Fehr
Prof. Dr. Jacob Goeree

The Faculty of Economics, Business Administration and Information Technology of the University of Zürich hereby authorizes the printing of this Doctoral Thesis, without thereby giving any opinion on the views contained therein.

Zürich, April 14, 2010

The Dean: Prof. Dr. Dr. Josef Falkinger

To my parents, with love.

Silent gratitude isn't very much to anyone.

---

attributed to Gertrude Stein

Acknowledgements

No researcher can ever work in isolation. The work I present in my dissertation is the fruit of extensive collaboration with many people, first and foremost with my supervisor Ernst Fehr who provided guidance and mentorship at every stage of my doctoral studies. Daria Knoch and Daniel Schunk worked with me on the study which I describe in chapter 1. As my "coach", Daniel was there for me anytime I needed advice. Experiment 3 is the result of intense exchange with my friend and fellow doctoral student Alain Cohn: of all people in the world, I could never have chosen a better partner in research.

Many other people were involved in the implementation of the two laboratory experiments: Urs Fischbacher whom I consulted with issues concerning his program z-Tree that I used to conduct the experiments, Stefan Schmid and the IT team who provided competent support, Franziska Heusi and Nina Spiri who managed the recruitment, Martin Hohmann who conducted the transcranial magnetic stimulation, and the many student assistants who helped with conducting the experimental sessions. The field experiment reported in experiment 3 was especially demanding precisely because field settings lack the infrastructure that is already in place at laboratories. Alain and I are therefore highly indebted to the industrious people at our cooperation partner in Berlin: I doubt that there are many companies out there that are as flexible and amenable to the needs of experimental economists as these people.

Experimental studies require much diligence on design, conducting, analysis, writing and—finally—publishing. During all these stages it is vital for scientific quality to put one's work in progress up for debate with colleagues. Because such constant exchange about projects is a necessary condition for

successful research, I am very grateful to conduct my doctoral studies at an institute that gives the opportunity to work with so many bright researchers. I can say that I have never encountered anyone here who did not have both an open ear and an open mind for all the issues with which I approached them. I would like to thank many members—past and present—of the Institute for Empirical Research in Economics, in particular: Björn Bartling, Alain Cohn, Charles Efferson, Lorenz Götte, Holger Herz, Michel Maréchal, Ryan McKay, Andreas Nicklisch, Daniel Schunk, Sonja Vogt and Christian Zehnder. Many other researchers gave valuable input to the work presented in this dissertation, among others George Akerlof, Stefano DellaVigna, Antonio Rangel and the participants at the Annual Research Forums of the NCCR "Affective Science" 2007 and 2009, at the World Meetings of the Economic Science Association 2007 and 2009, at the IZA Summer School 2007, at the ESE conference "Experiments in Organizations" 2009 and at smaller workshops and seminars.

The studies reported in chapter 1 and 2 are published, and I would like to thank the excellent editors, anonymous reviewers and proofreaders of Proceedings of the National Academy of Sciences and Proceedings of the Royal Society B who shaped the final stages of writing with their comments. I would also like to thank the Institute for Empirical Research in Economics, the University Research Priority Program "Foundations of Human Social Behavior" and the Swiss National Centre of Competence in Research (NCCR) "Affective Sciences" for their generous financial support.

Charles Efferson, Sally Gschwend and Ryan McKay were of invaluable help and great patience when correcting my English. Sally also translated the experimental instructions into English. As a notoriously disorganized person, I am heavily indebted to Helen Bernhard, Sally Gschwend and Nina Spiri who helped managing and structuring my work and schedule.

Finally, I want to express my deepest gratitude to the two persons to whom I owe everything: my parents. During all my life, you have given me so much unconditional love and infinite support. You made me what I am, and therefore this dissertation is ultimately your product.

Zürich, February 2010
Frédéric Schneider

Contents

List of Figures

List of Tables

> I grant that good and evil, *reward* and *punishment*, are the only motives to a rational creature: these are the spur and reins whereby all mankind are set on work, and guided.
>
> John Locke, Some thoughts concerning education (1693)

Introduction

This dissertation is composed of three chapters, each reporting an experimental study conducted during my doctorate. These three studies seem to be loosely connected at first glance, both in terms of content and of method. The most basic connections between the three chapters of this work are the general topic of "reciprocity" and the empirical approach of controlled experimentation. Indeed, the use of a wide variety of empirical methods is probably the distinctive feature of this dissertation. I would like to convince the reader that the logical connection between the three chapters is actually deep and that a broad methodological approach is not only helpful but in fact crucial for the understanding of human social behavior.

Chapter 1 describes how reputation concerns shape how much people cooperate with each other. Second movers in a repeated trust game situation with random matching behave much more cooperatively when their actions are observed by future first movers. This alone is nothing new and has been acknowledged by economists both theoretically (Kreps et al., 1982) and empirically (Andreoni and Miller, 1993). The fundamentally novel aspect of this experiment is the introduction of an additional treatment variable that taps into the tool box of neuroscience. By manipulating neurophysiological processes in our participants, we are able to provide causal evidence on the biological mechanism that makes people willing to cooperate strategically and why this behavior is characteristic for humans. As shown in experiment 1, the processing of immediate concerns on the one hand and strategic concerns on the other is anatomically differentiated in the brain. The prioritization

of these concerns and their realization in behavior is crucially related to the right dorsolateral prefrontal cortex (rDLPFC).

While chapter 1 describes cooperative behavior that is largely motivated by self-interested reasoning, it also hints at other motives because participants cooperate even in the absence of strategic incentives, another fact that experimental economists have reported time and again.[1] Some researchers (Haley and Fessler, 2005; Bateson et al., 2006) report suggestive evidence that even cooperation in anonymous interactions may also be the result of (unconscious) reputation motives which are evoked by "implicit reputation cues". Chapter 2 describes a laboratory experiment where the implicit reputation hypothesis for cooperation in the absence of strategic incentives is directly tested, again in a trust game. Cues were implemented in guise of stylized eye shapes on the background of the trustees' computer screens. The results suggest that implicit reputation motives are unlikely to drive cooperative behavior in anonymous interactions. In accord with chapter 1, this experiment reaffirms that *explicit* reputation incentives dramatically increase the amount of cooperation, although there is substantial reciprocity even in the absence of strategic concerns.

Many economists are comfortable with the idea that people privately care about fairness and reciprocity but insist that material incentives make them behave self-interestedly, and that market forces ultimately eliminate those people who do not comply with the rules of *homo oeconomicus*. They dismiss contradicting results from laboratory experiments because they assert that—in contrast to economic interactions outside the laboratory—the incentives in the lab are too low-powered to offset fairness concerns, that the laboratory setting is artificial, and that hence, results from lab studies are not readily generalizable (Levitt and List, 2007). Although this partial view is questionable, empirical researchers should strive for the widest array of evidence available, including evidence from both laboratory experiments and field experiments.

[1] See Fehr (2000) and Cooper and Kagel (2010) for reviews, and specifically Berg et al. (1995) for the seminal paper on trust games.

Chapter 3 of this dissertation therefore covers the economic relevance of fairness concerns outside of the laboratory in the case of employment relations. The empirical method of choice is the field experiment because it accommodates the concern that behavior in the laboratory may be artificial while retaining a maximum of control over the situational parameters. The employer-employee setting creates a "real life" version of a principal-agent social dilemma. Crucially, this study adds a further topical dimension to the research on reciprocal behavior: social comparison. Do workers compare themselves with others, and if so, does it influence their productivity? Experiment 3 reveals that people not only care about how they are treated, they also care about how they are treated *relative to others*. They reciprocate unequal treatment more negatively than equal treatment. This result helps explain important anomalies observed in empirical labor economics, namely wage compression and wage secrecy, and corroborates the fairness-based behavioral theory of efficiency wages and involuntary unemployment (Akerlof and Yellen, 1990).

Thus, the three experiments of this dissertation have a strong connection, both from a substantive and a methodological point of view. All three studies inquire into one particular behavior, namely second movers' contributions in a sequential social dilemma, as the natural measure of reciprocity. The research that I present here follows a clear line of reasoning in the exploration of human reciprocity and reaches from its evolutionary origins ("Why do people reciprocate?") and neurobiological causes ("What are the workings behind it?") to its relevance for economic research ("Why should economists care?").

Economists should care because many empirical phenomena can only be understood if economists open the black box of human decision-making. The guiding principles of human behavior, such as reciprocity, must be explored because they greatly influence economic outcomes. In a world where economic interactions are characterized by incomplete contracts, reciprocity plays a major role in keeping markets efficient. While concerns for reputa-

tion in repeated interactions can resolve the issue of incompleteness through strategic reciprocity (see experiment 1), concerns for fairness have the potential to make contracts "complete" even in anonymous one-shot interactions, through *strong* reciprocity (see experiment 2). On the other hand, strong concerns for fairness can also affect economic outcomes adversely, for example, when firms refrain from cutting wages in fear of negative worker reactions (see experiment 3). If actual wages remain above the market clearing wage, involuntary unemployment is the consequence.

By now, the scientific exploration of decision-making has become more powerful through the development of suitable empirical tools. It is one purpose of this thesis to show how some of these tools—"canonical" behavioral laboratory experimentation, neuroeconomic laboratory experimentation and field experimentation—can be used as complements for this purpose.

The findings of experiment 1 and experiment 2 are published in Knoch et al. (2009) and Fehr and Schneider (2009). The material presented in the appendix has been translated into English, documents in the original German versions as well as original z-Tree program files can be obtained from the author upon request.

Reputation is what you are in the light; character is what you are in the dark.

Chinese proverb

Experiment 1

The Neurobiological Foundation of Strategic Reciprocity

Summary Reputation formation pervades human social life. In this chapter, we investigate the neural underpinnings of this important social mechanism. We show that disruption of the right, but not the left, lateral prefrontal cortex (PFC) with low-frequency repetitive transcranial magnetic stimulation (rTMS) diminishes subjects' ability to build a favorable reputation. This effect occurs even though subjects' ability to behave altruistically remains intact when reputation incentives are absent and even though they are still able to recognize both the fairness standards mandating cooperation and the importance of cooperation for building a good reputation. Thus, subjects with a disrupted right lateral PFC no longer seem to be able to resist the temptation to defect, even though they know that this has detrimental effects on their future reputation. This suggests an important dissociation between the knowledge about one's own best interests and the ability to act accordingly in social contexts. These results link findings on the neural underpinnings of self-control and temptation with the study of human social behavior, and they may help explain why reputation formation remains less prominent in most other species with less developed prefrontal cortices.

1.1 Introduction

Humans are unique in the extent to which social norms regulate their lives, and reputation formation is a powerful mechanism in generating norm compliance. Although much norm compliance is voluntary, there is ample evidence that people are more likely to comply with norms when they feel observed by others. In such a situation—"in the light", as the Chinese proverb puts it—individuals signal their quality as cooperators to future interaction partners, thereby forming a good reputation.

Reputation formation is characterized by two features. First, the signals for building a good reputation (in human societies) have to be costly, otherwise they would be "cheap talk" and thus of no informational value for the potential interaction partner. Second, this process of costly reputation formation is characterized by a trade-off between the current benefits of defection and the future benefits of cooperation through a good reputation.

Evidence for the crucial role of a good reputation in social decision making comes from empirical studies showing that individuals increase their levels of cooperation and are more likely to comply with norms when they know that others observe their behavior, and that individuals cooperate with those whom they observe cooperating with others (Nowak and Sigmund, 2005; Milinski et al., 2002; Wedekind and Milinski, 2000; Nowak and Sigmund, 1998; Panchanathan and Boyd, 2004; Brown et al., 2004; Camerer and Weigelt, 1988; Fehr et al., 2009; Basu et al., 2009; Houser et al., 2006; Keser and Van Winden, 2000; Fehr and Gächter, 2000; Falk et al., 1999; Cochard et al., 2004; Engelmann and Fischbacher, 2009; Seinen and Schram, 2006). Thus, the concern for reputation profoundly affects our daily social interactions and motivates many important decisions in our lives.

Although reputation formation mechanisms are ubiquitous in social exchange, their neurobiological substrate remains largely unknown. Moreover, a universal question arises, one with relevance not only to cognitive neuroscience, but also to fields of research in evolutionary biology, developmental

psychology, and behavioral economics: Which skills are required to acquire a good reputation? Intuitively, we assume that there must be a self-control capacity because forming a reputation typically requires an individual to overcome the temptation to defect to gain future reputation benefits. From a neurobiological perspective, we thus assume the involvement of the PFC because this region has been shown to be involved in self-control processes (Aron et al., 2004; Miller and Cohen, 2001; Knoch et al., 2006).

Four previous neuroimaging studies have examined reputation (Takahashi et al., 2008; Delgado et al., 2005; Izuma et al., 2008; King-Casas et al., 2005). Two of these studies did not address the neural underpinnings of the process of individual reputation formation; i.e., they did not focus on the individual who forms a reputation. Instead, they examined individuals who made decisions based on reputation information about another individual (Takahashi et al., 2008; Delgado et al., 2005). In one of these studies, for example, subjects played iterative trust games with three partners whose (fictional) profiles make them seem morally good, bad, or neutral (Delgado et al., 2005). The study reveals that information about the moral reputation of the interaction partner affected the investors' reward prediction error signal in the caudate nucleus during reciprocal exchange. Another study showed activation of reward-related brain areas when a subject learned that others perceived his or her reputation as good (Izuma et al., 2008). Finally, one study used hyperscanning functional MRI (fMRI) to investigate two individuals interacting in an iterated trust game. The recorded brain activity showed that the peak activation of the caudate nucleus underwent a temporal shift as the reputation of the interaction partner developed (King-Casas et al., 2005).

No previous study provides causal evidence about the brain processes involved in costly reputation formation, however. Functional imaging methods, although indispensable, do not permit causal inferences about the effect of brain processes on human behavior because the observed neural activations could be spuriously correlated with task performance and need not necessar-

ily play a causal role in task execution (Walsh and Cowey, 2000; Sack and Linden, 2003). In contrast, brain stimulation techniques, such as transcranial magnetic stimulation (TMS), interfere non-invasively with the activity of defined areas in the human cortex, allowing researchers to observe the behavioral impact of an increase or decrease in the cortical excitability of the stimulated brain region. Application of low-frequency repetitive TMS (rTMS) for several minutes leads to suppression of activity in the stimulated brain region that outlasts the duration of the exposition to rTMS by about half the duration of the stimulation (Robertson et al., 2003; Eisenegger et al., 2008). In this chapter, we investigate the effect of disrupting the PFC by means of rTMS on subjects' reputation formation.

1.2 Experiment

1.2.1 Experimental Procedure

We studied 87 right-handed men (mean age 22.6 years, ranging from 20 to 27 years). All provided written informed consent to participate in the study, which was approved by the local Ethics Committee. None of the subjects had been previously subjected to TMS or a trust game. No subject had a history of psychiatric illness or neurologic disorder. There was no difference among the 6 experimental groups with respect to age ($\chi^2 = 8.478$; $df = 5$; $p = 0.1318$; Kruskal-Wallis test). None of the subjects experienced serious adverse side effects or reported scalp pain, neck pain, or headache after the experiment. The experiment was programmed and conducted with the z-Tree software (Fischbacher, 2007), recruitment was conducted using ORSEE software (Greiner, 2004).

We implemented a fully-crossed 2×3 factorial design with the two factors "information condition" and "stimulation condition" (see section 1.2.2). Subjects took part in only one of the six possible factor combinations (between-subjects design). Only trustees received stimulation. Because TMS could

be applied to only one subject at a time, investors came collectively to the laboratory of the Institute for Empirical Research in Economics ("investor sessions"), while trustees were located at the University Hospital of Zürich ("trustee sessions"). Implementation was such that investors and additional trustees completed the experiment in the computer lab before the TMS experiment. The subjects receiving stimulation then came individually to the University Hospital Zürich for the trustee sessions. In each period, each of these subjects was randomly matched with a subject from the investor sessions so that their histories were matched; for example, a trustee who had opted twice for "nothing" and once for "equalize" in the three previous periods would be matched with an investor who had observed the same play in the three previous periods in the investor session. The average net duration of the experiment (from onset of the trust game to completion of control questions) was 401.5 s (maximum, 478.8 s). All participants were paid according to their payoffs in the game; 1 point in the game equaled 0.20 Swiss francs (CHF).[1] In addition, participants in the lab received a show-up fee of CHF 10, and participants receiving stimulation received a show-up fee of CHF 60. We implemented the experiment in this way to ensure that subjects had a monetary incentive and a real concern about reputation.

1.2.2 Experimental Design

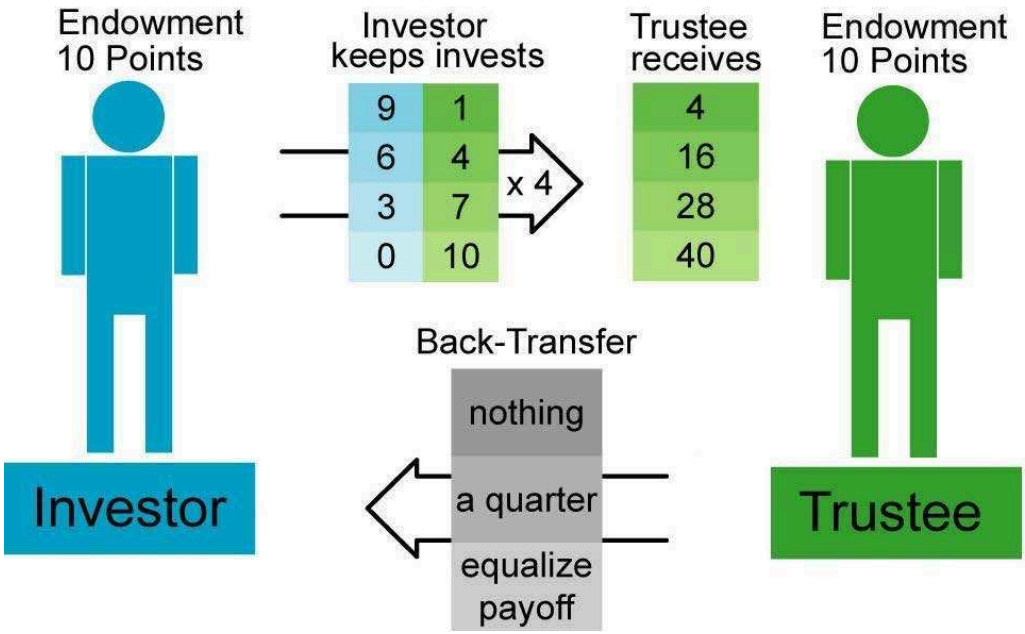
Trust Game

We chose a version of the trust game (Berg et al., 1995) as a vehicle for investigating the effects of rTMS on costly reputation formation (see Figure 1.1 for a schematic representation). Subjects played 15 periods of this trust game with randomly rematched partners each period. Each period was divided into 2 stages, an investor decision stage and a trustee decision stage.[2]

In the first stage, the investor was endowed with 10 points and had to decide the amount that he wished to invest in the current trustee. His choice

[1] CHF 0.20 = USD 0.18 at the time the experiment was conducted.
[2] Participant instructions are reproduced in appendix A.2.

Figure 1.1: Schematic representation of the trust game. In each period, two anonymous individuals—a first mover (investor) and a second mover (trustee)—receive an endowment of 10 points each. The investor can invest 1, 4, 7, or 10 points (e.g., investing 4 points means keeping 6 points). The experimenter quadruples the invested points and transfers them to the trustee, who then decides how many points he would like to transfer back to the investor. The trustee has three options, independent of the invested amount: he can transfer back nothing, a quarter of the received amount, or an amount that equalizes the period payoff between the investor and the trustee. Note that the game design is based on the design of experiment 2, which was conducted prior to the experiment presented in this chapter.



was restricted to 1, 4, 7, or 10 points.[3] Once the decision was made, the invested amount was quadrupled and passed to the trustee; i.e., the received amount was four times greater than the invested amount.[4] In the second

[3]Zero investments were not allowed in order to exclude "costless" reputation formation. Also note that no exact medium investment was allowed; i.e., investors were forced to choose either a high or a low investment.

[4]After pretesting, we decided to quadruple the amount—instead of the usual tripling—

stage, the trustee then had to decide how much of the quadrupled amount he wanted to transfer back to the investor. This back transfer was not quadrupled.

The trustee's choice was restricted to the following:

(a) Back-transfer nothing (= 0 percent of the received amount)
(b) Back-transfer 25 percent of the received amount (= the invested amount). The investor finishes the period with his endowment of 10 points, allowing him to "break even."
(c) Transfer an amount that equalizes payoffs between the 2 subjects (= 62.5 percent of the received amount).

The restriction to three trustee choices also has the added advantage that the reputation implications of different trustee behaviors are transparent. For example, paying back nothing is unambiguously bad for the formation of a good reputation, while equalizing payoffs is unambiguously good. The payoff structure of this trust game is shown in Table 1.1.[5]

Table 1.1: Payoff matrix for investor and trustee. The investor plays rows (investment of 1, 4, 7 or 10 points), the trustee plays columns (back transfer is nothing, compensate or equalize). The first number in a cell is the investor payoff, the second number the trustee payoff.

	Trustee					
Investor	nothing		compensate		equalize	
1 point	9	14	10	13	11.5	11.5
4 points	6	26	10	22	16	16
7 points	3	38	10	31	20.5	20.5
10 points	0	50	10	40	25	25

to ensure sufficient 10-point investments in the anonymous condition, allowing for profound statistical inference.

[5]Note that this structure is identical to the one in experiment 2.

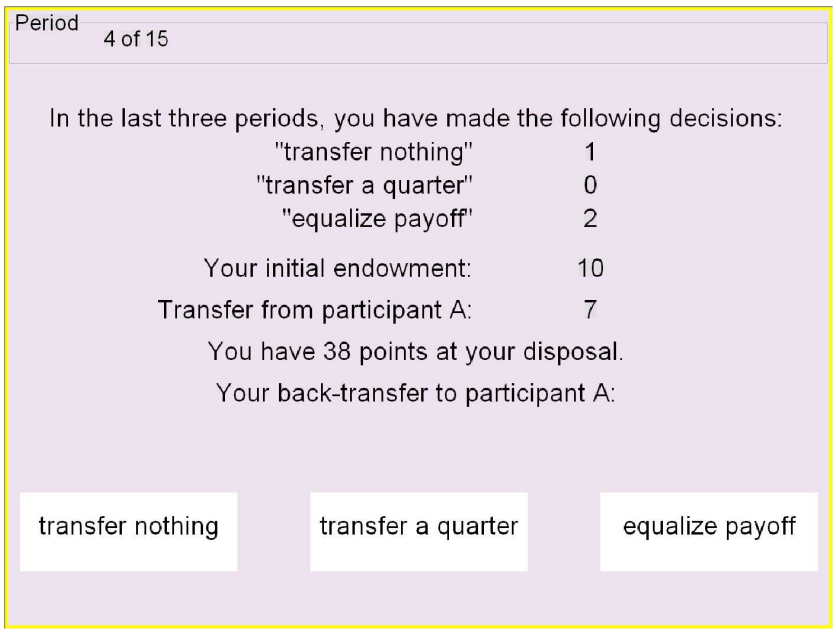
Information Conditions

We implemented two information conditions, which we term "anonymous condition" and "reputation condition." In the anonymous condition, the trustee's previous decisions are unknown to the current investor. In the reputation condition, the investor has some information about the trustee (see Figure 1.2). In particular, he can observe the trustee's decisions in the previous three periods (i.e., how many times the trustee chose to transfer back "nothing," "a quarter," or "equalize payoffs").[6] Thus, the trustee knows that the investor can condition his transfer on the trustee's previous three back transfer decisions. This also means that the trustee's back transfer in the current period affects the information that future investors receive about him, i.e., it affects his reputation in future periods. A trustee is therefore likely to acquire a bad reputation by transferring back "nothing," whereas a trustee improves his reputation by choosing to "equalize payoffs." Because a trustee who transfers nothing is unlikely to receive high transfers from the investors in future periods, the trustees have an incentive to make high back transfers in the reputation condition. Thus, this reputation incentive generates a motivational conflict for the trustees. A trustee could maximize his short-run self-interest by choosing to transfer nothing back to the investor in the current period, but this action is likely to have detrimental effects for his reputation and decreases future investors' willingness to transfer money to him. Therefore, to reap the benefits from a good reputation in future periods, a trustee must constrain his immediate self-interest and forgo the current option of transferring back nothing.

In the reputation condition, each of the trustee's decisions (given past decisions) had exactly the same relevance for his reputation, thus keeping the strategic incentive associated with the choice constant. This was achieved by three design features: a stable information window of three periods, no

[6]Investors could only observe previous trustee choices, not back transfers. They had no information about the size of the corresponding investments or the chronological order of the choices.

Figure 1.2: Trustee decision screen. On the top section of the screen, the trustee can see his decisions in the three previous periods (not in chronological order). In the reputation condition, the trustee also knows that the investor is informed about the trustee's previous three decisions before he makes a transfer decision. The middle section of the screen contains information about the current investor's transfer and the resulting points at the trustee's disposal. The bottom section features three clickable buttons for the trustee's decision.



information on investment levels, and no information about the sequence of choices. Moreover, these features helped keep the cognitive effort of the game low. While the reputation benefit of a particular decision was kept constant, note that the *immediate* benefit of the trustee's choice was proportional to the amount that the investor transferred (see Table 1.1). This made it possible to observe different levels of "temptation."

In contrast, the strategic incentive for behaving in a cooperative manner is completely absent in the anonymous condition because the investors have no information about the trustees' past behavior. In terms of the Chinese proverb cited at the beginning of this chapter, the trustees in the anonymous

condition act in perfect darkness, and only their "character" plays a role. Thus, the anonymous condition measures how much a trustee is willing to return voluntarily to the investor (which may be viewed as a form of altruistic behavior). This amount reflects the trustee's preference for back transfers if there are no strategic reputation incentives. If the trustee returns amount x to the investor in the anonymous condition, then the trustee is apparently not willing to return more than x. However, the trustee very well might return more than x in the reputation condition, as strategic incentives for reputation formation are then present; i.e., in his case the trustee must override his immediate self-interest to build a good reputation.

Stimulation Conditions

We applied low-frequency rTMS for 15 min to 87 healthy subjects acting in the role of the trustee.[7] To investigate a possible hemispheric laterality in the role of lateral PFC on trustee decisions, we applied rTMS to the right DLPFC or to the left DLPFC. The creation of a stimulation group receiving rTMS to the right DLPFC and a control group receiving rTMS to the left DLPFC was important to control for the potential side effects of rTMS (Abler et al., 2005), including discomfort, irritation, and mood changes. We implemented an additional control condition in which we applied sham stimulation for 15 min to the right or left DLPFC. Thus, the factor "stimulation" had three conditions: right DLPFC, left DLPFC, and sham. As described in section 1.2.2, we implemented an anonymous condition and a reputation condition. Thus, the experiment had a 2×3 design, with the factors "information" (anonymous, reputation) and "stimulation" (left rTMS, right rTMS, sham) leading to six experimental groups. We randomly assigned each subject to one of the six groups.

[7] See appendix A.1 for technical specifications.

1.2.3 Questionnaire

Because disruption of the PFC also might affect subjects' perception of what constitutes the social norm in a certain situation, we further elicited individuals' perception of fairness norms immediately after the trust game by confronting them with a hypothetical scenario. We asked participants to judge the fairness of a hypothetical trustee's behavior on a 7-point scale from "very unfair" to "very fair." The scenario described an investor who invests 7 points while the trustee returns nothing.

Disruption of the PFC also might affect subjects' ability to assess the consequences of a particular reputation, i.e., to assess the impact of actions on future social interaction, which is an abstract and cognitively demanding task. To rule out this explanation, we used another scenario to measure an individual's assessment of the potential consequences of a certain reputation. We asked the subjects how many points (1, 4, 7, or 10) they would expect an investor to transfer to a trustee who had chosen to "equalize payoffs" twice and to transfer back nothing once.

Subjects also completed personality questionnaires that assessed impulsivity (Carver and White, 1994), using the BIS and BAS scales, and personal norms of reciprocity (Perugini et al., 2003). These questionnaires were completed roughly 10 days after the experiment. Details regarding the questionnaire are provided in appendix A.3.

1.3 Hypotheses

How will disruption of the PFC with low-frequency rTMS affect the trustees' behavior? Because the lateral PFC has been shown to be reliably involved in overriding prepotent responses and self-control processes (Aron et al., 2004; Miller and Cohen, 2001; Knoch et al., 2006), and because costly reputation formation requires overriding immediate benefits, disrupting this brain region should functionally weaken self-control capacity and thus lead to a

lower back transfer in the reputation condition compared with the other stimulation groups. In contrast, little or no self-control effort is involved in the anonymous condition because the trustee has no reputation incentive to transfer back more than his immediate preference dictates. Therefore, we would expect to see little difference between the stimulation groups for the anonymous condition. Moreover, because the right lateral PFC in particular has been shown to be involved in control capacities (Aron et al., 2004; Knoch et al., 2006), we hypothesized that disruption of the right, but not the left, lateral PFC will lead to difficulty resisting the temptation to go for the immediate benefit and thus reduce the ability to form a good reputation.

It is important to note that the trustee in the reputation condition knows that the investor has information only about his three previous choices (i.e., "nothing," "a quarter," or "equalize payoffs"), not about how high the corresponding investments were in the previous three periods. For example, if a trustee receives an investment of 1 point and chooses "equalize payoffs" to form a good reputation, then he actually transfers 2.5 points back. This is because he received 4 points (1 point, quadrupled by the experimenter), and transferring back 2.5 of those 4 points together with the initial endowment of 10 points leaves both the investor and the trustee with equal amounts totaling 11.5 points. If a trustee receives an investment of 10 points and chooses "equalize payoffs," then his back transfer is 25 points, and both players end up with a total of 25 points. Because future investors will observe only the choice "equalize payoffs" but not the amount actually transferred by the investor, the trustee's reputation benefit is the same in both cases. The immediate costs, however, are different: 2.5 points in the first case and 25 points in the second case.

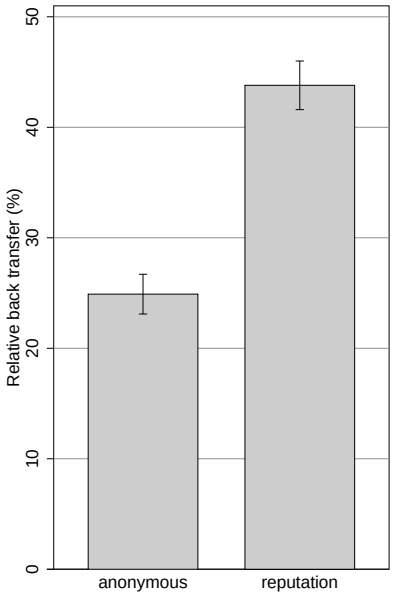
Therefore, the costs of reputation formation (i.e., the number of points the trustee must forgo to form a good reputation) vary with the size of the investment, while the effect of a particular choice on a trustee's reputation is always the same, regardless of the received investment. In other words, while the future reputation value of a trustee's choice is independent

of investments, the immediate cost of reputation formation, and thus the temptation to maximize one's short-run self-interest, varies with the size of the investment. Thus, the self-control effort necessary to constrain short-run self-interest is likely to be much higher in cases of a large investment compared with a small investment, where reputation formation is almost costless. This variation in the temptation to maximize one's short-run self-interest by paying back nothing enables us to investigate whether the effect of disrupting the lateral PFC depends on the degree of self-control required for reputation formation. While the self-control hypothesis predicts a specific effect for high investments, a *global* effect of rTMS on back transfers—irrelevant of the invested amount—would not be consistent with this hypothesis. Rather, such an effect would suggest other explanations, such as diminished ability to to take into account the future consequences of one's own decisions.

1.4 Results

Our results show that reputation formation paid off for the trustees in the long run because investors gave more points to trustees who cooperated in the past than to defectors. Trustees had a 71 percent probability of receiving a 10-point investment if they always equalized payoffs, dropping below 6 percent if they always chose to transfer back nothing. Consequently, a strategy of cooperating in the first 14 periods and defecting in the last period (i.e., rational cooperation) was on average 43 percent more profitable (371 points) than always defecting (260 points). Thus, the trustees had an incentive to constrain their short-run self-interest and to transfer back a high amount in the reputation condition because the investors conditioned their investments on the trustee's past actions. Accordingly, our results show that trustees cared greatly about their reputation when reputation formation was possible. Subjects sent back on average 24.9 percent of the transferred amount in the anonymous condition, compared with 43.8 percent in the reputation condition (Figure 1.3).

Figure 1.3: Effect of information condition on back transfers. Error bars represent standard errors of the mean (clustered over subjects), observations are pooled over all stimulation conditions.

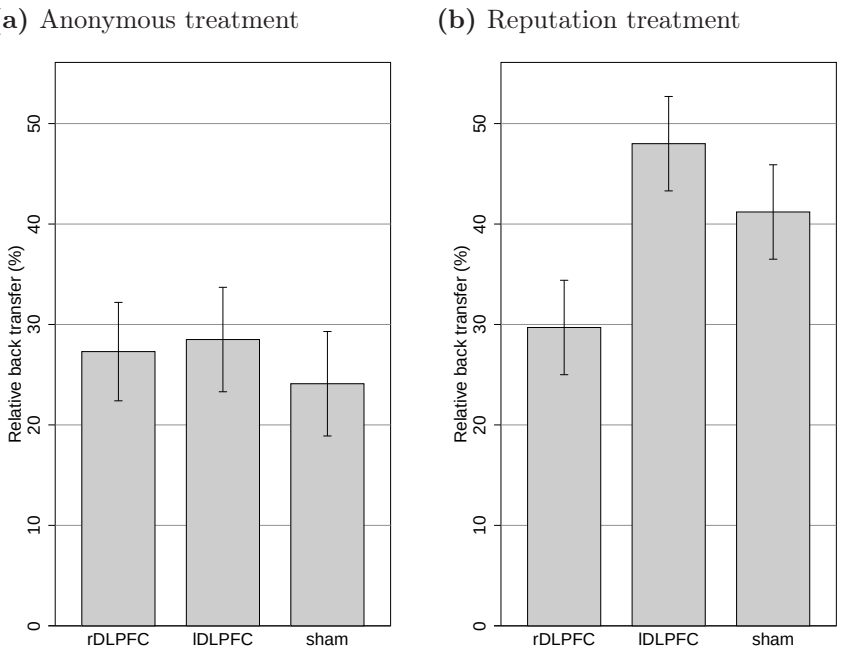


Of primary interest are back transfers with regard to the investors' highest investment because the temptation to follow short-run self-interest, and thus the requirement for self-control effort, is greatest in this case. Focusing on the reputation condition (Figure 1.4b), we see that the back transfer for the highest investment was 41.2 percent following sham rTMS and 48.0 percent after real rTMS of the left DLPFC.[8] These results contrast sharply with the back transfer of 29.7 percent after rTMS of the right DLPFC. The differences in back transfers across the stimulation groups are significant in the reputation condition ($p < 0.001$ for the difference between right and left DLPFC and $p = 0.015$ for the difference between right DLPFC and sham condition). In contrast, we found no significant differences in back transfers among the three stimulation groups in the anonymous condition (Figure 1.4a,

[8]This section reports the results of a generalized least squares (GLS) regression. For details, see appendix A.4.

$p = 0.816$ for the difference between right and left DLPFC and $p = 0.232$ for the difference between right DLPFC and sham).

Figure 1.4: Trustee's mean back transfer after an investment of 10 points, across stimulation conditions. In the reputation condition, subjects whose right DLPFC was disrupted transferred back significantly less points than those in the other 2 stimulation groups ($p < 0.02$). Error bars represent standard errors of the mean, §clustered over subjects.



In other words, while disruption of the right DLPFC significantly reduced back transfers in the reputation condition in cases of highest investment, it did not reduce back transfers in the anonymous condition. This indicates a significant differential effect of rTMS across stimulations (right DLPFC versus left DLPFC and sham) in the reputation condition, but not in the anonymous condition.

Interestingly, those subjects in the reputation condition who received rTMS to the right DLPFC transferred similar amounts back to the investor as those in the anonymous condition (compare Figure 1.4a and 1.4b; $p = 0.667$;

t-test). Thus, disrupting the right DLPFC completely removed the behavioral impact of the reputation condition, but had no effect on behavior when reputation formation was not possible. Moreover, there were no significant differences across stimulation groups for lower investments, where the temptation reap short-run gains and thus the recruitment of self-control effort was lower (all $p > 0.193$).

rTMS of the right DLPFC limited subjects' ability to override immediate short-run benefits. However, rTMS changed neither subjects' perception of the fairness norm nor their ability to assess the consequences of past and current trustee behaviors on future investments, which we elicited immediately after the experiment (see appendix A.3). First, subjects in all three stimulation groups judged the scenario of transferring back nothing in response to an investment of 7 as rather unfair, and there were no differences in fairness judgments across groups ($p = 0.376$; Kruskal-Wallis test). Second, rTMS of the right DLPFC did not change subjects' ability to assess the consequences of past and current trustee behaviors because subjects in the different stimulation groups predicted the same investments by future investors in response to a given profile of past back transfers ($p = 0.950$; Kruskal-Wallis test). Moreover, if rTMS of the right DLPFC had impaired subjects' general ability to perform complex calculations, then we would have observed differences across stimulation groups for the lower investments as well; however, our results show a behavioral effect only for the highest investments. This indicates that disruption of the right DLPFC had an effect on the behavioral ability to form a good reputation, even though it did not affect subjects' ability to perform complex cognitive operations, their recognition of the prevailing fairness norm, or their ability to assess the future consequences of back transfer behaviors.

We also investigated whether individual differences in impulsivity and the propensity to reciprocate kind or hostile acts can explain our results. We found that neither dispositional differences in subjects' reciprocity norm nor individual differences in impulsivity across treatment groups can explain

the behavioral differences across conditions; there was no difference across treatments for impulsivity or for the reciprocity norm.[9]

1.5 Conclusion

Our results indicate a highly specific, lateralized effect of a disrupted function of the lateral PFC on the ability to form a reputation for being trustworthy. We found no differences between the stimulation groups in the anonymous condition, where the incentives for reputation formation are absent. In this condition, only subjects' preferences for altruistic behaviors can induce them to repay trust, implying that an interference with the function of the right lateral PFC leaves their altruistic propensities to behave in a trustworthy manner unchanged. This contrasts with an rTMS effect in those circumstances in which costly reputation formation requires a particularly strong recruitment of self-control effort, i.e., when the investors make a high investment. In this situation, the incentive to yield to the short-run costs for building a reputation is greatest, suggesting an interpretation of the rTMS effect in terms of the reduced ability to recruit the required self-control resources. The absence of any rTMS effect on subjects' ability to recognize the prevailing fairness norm supports this conjecture. Thus, despite the fact that subjects are well aware of the existing fairness norm, and even though they have pecuniary incentives to obey this norm in the reputation condition, they nevertheless do not act accordingly. This suggests that rTMS causes a specific inability to constrain short-run temptations, rather than a cognitive inability to perceive the normative demands involved in the situation. The finding that rTMS had no effect on subjects' ability to assess the future consequences of past back transfers further supports our interpretation. Subjects across all three stimulation conditions had the same knowledge about the future benefits of high current back transfers, but only those subjects

[9]Behavioral Inhibition System (BIS) scale: $p = 0.827$; Behavioral Approach System (BAS) scale: $p = 0.967$; positive reciprocity scale: $p = 0.741$; negative reciprocity scale: $p = 0.971$; Kruskal-Wallis tests.

with transiently disrupted right DLPFC function were less able to constrain their short-run self-interest and thus exploit this knowledge.

Taken together, these results support the hypothesis that right, but not left, lateral PFC activity is a crucial factor in the ability to forgo immediate benefits to form a good reputation. Subjects whose right lateral PFC was disrupted behaved as if they were not concerned about their reputation when reputation formation required forgoing a large current benefit, suggesting that they were less able to pay an immediate cost for future social reputation benefits even though their ability to assess these benefits cognitively remained intact. These findings suggest an important dissociation at the neurobiological level between the knowledge about what is in one's own best interest in social interaction situations and the ability to act accordingly. Moreover, by providing causal evidence on the role of the prefrontal cortex in costly reputation formation, our findings also may help explain why reputation mechanisms are rare in other species with less developed prefrontal regions.

In highly complex processes such as reputation formation, brain areas do not act in isolation, but rather must work together as a network. Future studies could combine low-frequency rTMS and fMRI to explore how different brain regions interact on the functional anatomical level in reputation formation.

Two things fill the mind with ever new and increasing admiration and awe, the more often and steadily reflection is occupied with them:
the starry heaven above me and the moral law within me.

Immanuel Kant, Critique of Practical Reason (1788)

Experiment 2

The Evolutionary Mechanism behind Strong Reciprocity

Summary Strong reciprocity is characterized by the willingness to altruistically reward cooperative acts and to altruistically punish norm-violating, defecting behavior. Recent evidence suggests that subtle reputation cues, such as eyes staring at subjects during their choices, may enhance prosocial behavior. Thus, in principle, strong reciprocity could also be affected by eye cues. In this chapter, we investigate the impact of eye cues on trustees' altruistic behavior in a trust game and find zero effect. Neither the subjects who are classified as prosocial nor the subjects who are classified as selfish respond to these cues. In sharp contrast to the irrelevance of subtle reputation cues for strong reciprocity, we find a large effect of explicit, pecuniary reputation incentives on the trustees' cooperativeness. Trustees who can acquire a good reputation that benefits them in future interactions honor trust much more than trustees who cannot build a good reputation. These results cast doubt on hypotheses suggesting that strong reciprocity is easily malleable by implicit reputation cues not backed by explicit reputation incentives.

2.1 Introduction

As discussed in chapter 1, reputation incentives increase people's propensity to behave cooperatively. A large literature confirms that human cooperation is positively affected by the possibility of acquiring a 'good' reputation that may pay off in future interactions (Gächter and Fehr, 1999; Wedekind and Milinski, 2000; Milinski et al., 2001, 2002; Brown et al., 2004; Rege and Telle, 2004; Rockenbach and Milinski, 2006; Kurzban et al., 2007; Engelmann and Fischbacher, 2009; Fehr et al., 2009). However, the fact that reputation incentives improve prosocial behavior does not mean that people do not cooperate in the absence of such incentives. In experiment 1, e.g., trustees in the anonymous condition transferred back on average 25 percent of the amount they received. These back transfers can be interpreted as altruistic.

Human altruism represents a huge outlier in the animal world (Boyd and Richerson, 2005). Humans often behave altruistically towards genetically unrelated strangers, even if the chance of meeting these strangers again is extremely small and reputation concerns are unlikely to play a role (e.g., tipping an unknown taxi driver in a large foreign city). Altruistic behavior in the absence of any opportunity of repeated interaction and reputation formation has been repeatedly shown in tightly controlled economic experiments (Camerer, 2003; Fehr and Fischbacher, 2003; Gintis et al., 2003). Experimental evidence (Fehr et al., 2002), social preference theories (Rabin, 1993; Fehr and Schmidt, 1999; Falk and Fischbacher, 2006; Dufwenberg and Kirchsteiger, 2004), and evolutionary theories (Gintis, 2000; Henrich and Boyd, 2001; Boyd et al., 2003; Bowles and Gintis, 2004) also indicate that a special type of altruistic behavior—strong reciprocity—plays a particularly important role in establishing and sustaining cooperation among strangers. Strong reciprocity is characterized by the willingness to altruistically reward cooperative acts and to altruistically punish norm-violating, defecting behavior. As a consequence, strong reciprocity generates important incentives for cooperation among strangers.

Recent articles even seem to suggest that much of human altruistic be-

havior may be related to reputation motives. They argue that cooperation in anonymous one-shot interactions may be merely a response to *subtle* reputation cues that are not in fact related to the possibility of benefiting in future interactions from current altruistic acts (Haley and Fessler, 2005; Bateson et al., 2006; Hagen and Hammerstein, 2006; Burnham and Hare, 2007; Rigdon et al., 2009). Haley and Fessler assert that reputation incentives in the ancestral evolutionary environment thoroughly molded human social interactions because "natural selection can be expected to have shaped human psychology to be exquisitely sensitive to cues that are (or were, under ancestral conditions) informative with respect to the likely profitability of co-operation in a given situation" (Haley and Fessler, 2005, p. 249). These authors thus administered a visual cue—eyes staring at the subjects during decision-making—in an anonymous experimental game, a cue "that, over the course of human evolution, would have reliably indicated the potential observability of one's behavior" (p. 249). Haley and Fessler (2005); Bateson et al. (2006); Burnham and Hare (2007); Rigdon et al. (2009) indeed found that eyes staring at the subjects cause an increase in prosocial behavior in anonymous games such as the dictator game.

In this chapter we examine whether a reputation cue like that implemented in Haley and Fessler (2005) also affects strong reciprocity, by conducting an anonymous one-shot trust game in which a investor can send money to a trustee; the experimenter then quadruples this amount, so that the trustee receives four times the amount sent. The trustee observes how much the investor has sent and can then send back as little or as much money as he wants. Thus, the trustee can altruistically reward investors who have sent money, which constitutes an instance of strong reciprocity. By comparing the eye cue condition with a baseline condition without such cues we can assess the impact of eye cues on strong reciprocity.

In addition to the eye cue condition we implement another reputation condition in which subjects face a real pecuniary incentive for acquiring a good reputation. Previous work has argued that eye cues activate reputa-

tion concerns, but has not explicitly compared the effect of eye cues with the effect of explicit pecuniary reputation incentives. If humans are indeed "exquisitely sensitive" to reputation cues even if they carry no real pecuniary incentive power, eye cues should generate behavioral patterns that resemble those generated by explicit pecuniary reputation incentives. Our design enables us to draw this comparison and investigate the relative importance of reputation cues for altruistic behavior.

We also go beyond previous work by examining which—if any—subjects respond to the implicit reputation cue, because we measure subjects' degree of selfishness and opportunism with a Machiavellianism questionnaire (Christie and Geis, 1970). Assessing individual differences in subjects' responses to reputation cues is important because on average one might find a null effect that hides important inter-individual differences. Implicit reputation cues could increase the altruistic behavior of prosocial subjects, i.e., those who score low on the Machiavellianism (Mach) scale. This has important consequences on the interpretation of altruistic behavior in anonymous one-shot experiments. If prosocial subjects primarily respond to the implicit reputation cue, it is possible to argue that they are mostly prone to all kinds of other subtle reputation cues that are often not controlled for by the experimenter in the typical laboratory experiment (e.g., the mere presence of other subjects and the experimenter in the room, or simply hearing human voices). It would then be more plausible to attribute the baseline altruism observed in anonymous one-shot experiments to such uncontrolled reputation cues. However, if prosocial subjects do not respond to the eye cues, it is hard to argue along these lines. It is then implausible to attribute the observed altruistic behavior to uncontrolled subtle reputation cues. Thus, by measuring subjects' Mach scores we can put important constraints on the interpretation of altruistic behavior in anonymous one-shot experiments.

2.2 Experiment

2.2.1 Experimental Procedure

We conducted the experiment at the laboratory of the Institute for Empirical Research in Economics at the University of Zürich.[1] Participants were mostly students from the University of Zürich and the Swiss Federal Institute of Technology in Zürich. After arrival at the laboratory, participants were seated in separated compartments in front of computer terminals. Before the experiment started, participants had to answer five control questions about the experiment to make sure that they understood the instructions. The experimenter left the room after having checked the answers of all participants.[2]

We measured strong reciprocity as second-mover behavior in a series of one-shot trust games. A investor and a trustee interact with each other in a trust game. The investor can send money to the trustee, which is then multiplied by the experimenter so that the overall money available to the two parties increases. The trustee can then send back none or some of the money to the investor. Fairness norms typically demand that the trustee sends back some of the money he received, but the trustee is completely free to send back nothing if he likes. Details of the trust game are described in subsection 2.2.2.

Our experimental design includes three treatments: a baseline treatment where the trustee faces a neutral background screen (see Figure 2.1a); an "implicit reputation" cue treatment where the background screen features eyespot-like shapes (see Figure 2.1b), similar to those in Haley and Fessler (2005)[3]; and an "explicit reputation" treatment where the current investor is

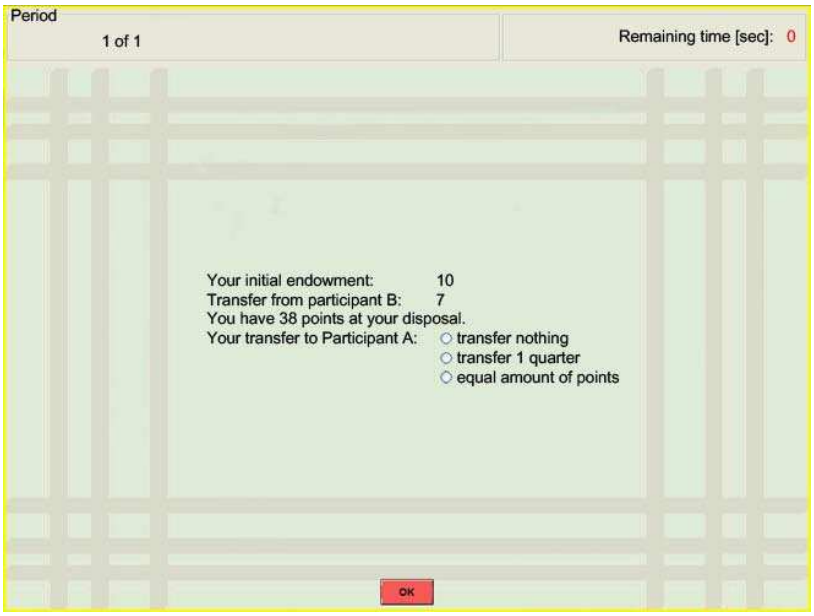
[1] The experiment was programmed and conducted with the z-Tree software (Fischbacher, 2007), recruitment was conducted using ORSEE software (Greiner, 2004).

[2] Translated versions of the instructions are reproduced in appendix B.

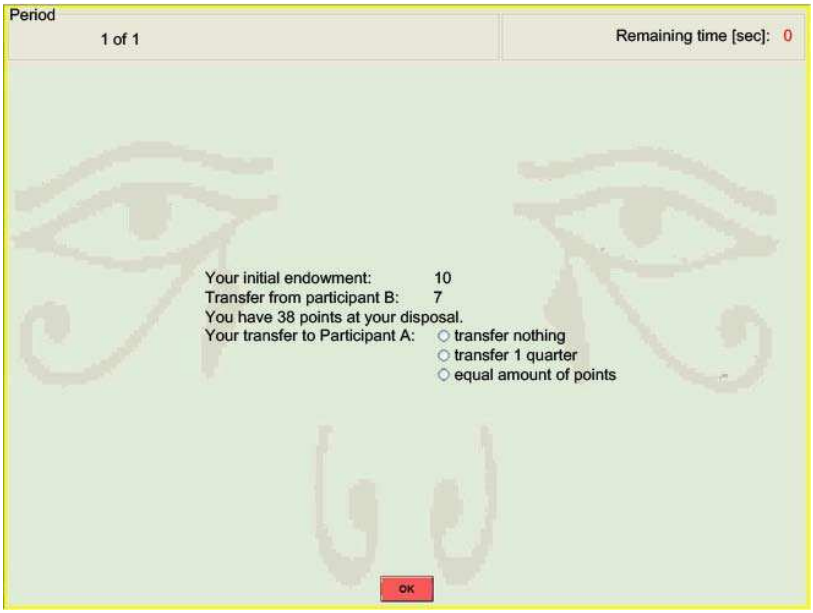
[3] We used the shape of Haley and Fessler's eyespot cues and matched the neutral and eyespot backgrounds for luminance and contrast. The background color was matched with z-Tree's default background color.

Figure 2.1: Trustee decision screen.

(a) Baseline background.



(b) Eyespots background.



informed of the trustee's decisions in the previous periods (same background screen as in the baseline treatment). Subjects were seated in separated compartments and were assigned either to the role of a investor or a trustee. They maintained their roles throughout the experiment. They then played 10 periods of the trust game, with randomly rematched partners each period.

We conducted eight sessions—three sessions each in baseline and implicit, two in explicit treatment—with 288 participants in total (144 investors and 144 trustees). The decisions in a group of subjects who interact with each other over the 10 periods are statistically not independent. In order to establish statistically independent observations, we created three matching groups per session, each consisting of 12 subjects. Only the subjects within a matching group were matched with each other during the experiment, generating three independent observations per session. With three matching groups per session, we have nine independent observations both in the baseline treatment and the implicit cues treatment, and six independent observations in the explicit reputation treatment.

Immediately after the end of the last period, the participants had to fill out a questionnaire containing items on emotional state, fairness attitudes, Machiavellianism, trust, and socioeconomic data. After completion, participants were paid a show-up fee of CHF 10 plus their earnings from the experiment, at the rate of 1 point = CHF 0.2. In total, a session lasted approximately 2 h and subjects earned on average CHF 48.88 (= USD 41.77).

2.2.2 Experimental Design

Each period of the experiment was a one-shot trust game. At the beginning of each period investors and trustees were endowed with 10 points. The game itself consisted of two stages: an investment stage, where investors had to decide how many points they would transfer to their current trustee, and a back transfer stage, where trustees had to decide how much they wanted to transfer back to the investor. The amount investors sent was quadrupled

and transferred to the trustee. Investors could choose between four possible transfers: 1 point, 4 points, 7 points, or 10 points.[4] Trustees had three options: they could transfer back either nothing, or the amount the investor sent (henceforth "compensate"), or they could transfer back an amount that equalized the period payoff between investor and trustee (henceforth "equalize"). When the trustee determined the back transfer he was perfectly informed about the investor's choice and thus did not have to form beliefs about the size of the investment.

We discretized the trustees' strategy space because this made it much easier to inform the investors about the trustees' past behavior in the explicit reputation treatment. Because the trustees had only 3 choices—"nothing", "compensate" and "equalize"[5]—we could present the information about a trustee's past behavior by informing the current investor how often the trustee had chosen each of these choice options in the preceding periods. In addition, the choice options "nothing," "compensate," and "equalize" are clearly distinct from each other, making it easier for the subjects to assess their consequences in terms of reputation.

Table 2.1 shows the payoff matrix that corresponds to our trust game. The first number in each cell of the matrix represents the investor's payoff, the second number denotes the trustee's payoff. For any given investment level, the trustee is always best off in terms of monetary payoff by transferring back nothing. This means that positive back transfers (i.e., the choices "compensate" and "equalize") can be regarded as altruistic acts because the trustee gives up some of his own payoff to increase the investor's payoff.[6]

Investors in the Explicit treatment could see a table with information about the current trustee's past behavior. This table contained the whole history of choices, however, not in chronological order (see instructions in

[4]Note that all four choices are positive. We did not include a "zero investment" choice because in this case a trustee could have built a reputation at no cost by choosing "equal split" in response to a zero investment in the reputation treatment.

[5]These terms were not used in the instructions.

[6]The trust game is identical to the trust game in experiment 1. There, however, investors had only information about the three most recent decisions of the current trustee.

appendix B.1.2 for an example). Note that the actual amounts transferred cannot be inferred from choices.

Table 2.1: Payoff matrix for investor and trustee. The investor plays rows (investment of 1, 4, 7 or 10 points), the trustee plays columns (backtransfer is nothing, compensate or equalize). The first number in a cell is the investor payoff, the second number the trustee payoff.

	Trustee					
Investor	nothing		compensate		equalize	
1 point	9	14	10	13	11.5	11.5
4 points	6	26	10	22	16	16
7 points	3	38	10	31	20.5	20.5
10 points	0	50	10	40	25	25

2.2.3 Questionnaire

As the response to the different treatments may be heterogeneous depending on the subject's degree of selfishness, we also measured each subject's Mach score. For this purpose we used the MACH-IV Machiavellianism Questionnaire (Christie and Geis, 1970), which provides a measure of selfishness and opportunism. Recent results from a neuroeconomic study (Spitzer et al., 2007) indicate that Machiavellian subjects are much less willing to share money in a dictator game, and respond much more strongly to pecuniary punishment threats for norm violations. Moreover, subjects' Machiavellianism also correlated strongly with activation in the lateral orbitofrontal cortex, known to be reliably activated when subjects face punishing stimuli. Thus, behavioral and neurophysiological evidence suggests that subjects' Machiavellianism may affect their responses to our treatment conditions.

We also measured subjects' fairness standards by asking them the following question: "Suppose that participant A [i.e., the investor] transferred 10 points to participant B [i.e., the trustee]. B then chose 'compensate'. How

fairly do you judge this behavior?"[7] Subjects indicated their answer to this question on a Likert scale, coded from 1 ('very unfair') to 7 ('very fair'). Note that subjects with high fairness standards perceive the choice as unfair and therefore assign a low score to this question, while subjects with low fairness standards perceive the choice as fair and assign a high score. We wanted the subjects to judge the action "compensate" because here the ambiguity about the fairness of the action is greatest, thus maximizing the chances for meaningful heterogeneity in subjects' responses.

2.3 Hypotheses

The implicit cues treatment measures the impact of implicit reputation cues on trustees' altruistic behavior in the trust game. The explicit reputation treatment enables us to assess the effect of explicit pecuniary reputation incentives on trustees' behavior. Thus, we can gain insight into the relative importance of the two kinds of reputation effects by comparing the effect of implicit cues with the effect of explicit reputation incentives.

Consider the baseline and the implicit reputation treatment condition first. The game played in these two conditions constitutes a true one-shot game because the players remain fully anonymous and they meet a different anonymous partner in each period. Therefore, if both players are completely selfish and want to maximize their money earnings, and the investor knows this, the following outcome is predicted. The selfish trustee will always choose "nothing" (i.e., his back transfer is zero), and the investor will invest the lowest possible amount because he knows that the trustee will transfer back nothing in any case.

However, there is a large literature indicating that a substantial share of experimental subjects is not completely selfish (see Fehr and Fischbacher (2003) for a review). This literature indicates that subjects may also have social motives such as inequity aversion (Fehr and Schmidt, 1999; Dawes

[7] Additions in square brackets did not appear in the questionnaire.

et al., 2007) or intention-based reciprocity (Rabin, 1993; Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006). Inequity-averse trustees will choose the "equalize" option, while trustees who interpret high investments as particularly kind acts will make more generous back transfers in response to high investments. We summarize both these behaviors under the term "altruistic rewarding" because they imply a benefit for the investor at the expense of the trustee and they reward the investor's transfer.

A key question then is whether subjects' social preferences are affected by implicit reputation cues such as eye spots. Recent evidence (Haley and Fessler, 2005; Bateson et al., 2006; Burnham and Hare, 2007) suggests that eye cues affect prosocial behavior in dictator games and public good games. In view of this literature, one would expect the trustees to respond to the eye spots in the implicit reputation treatment by making significantly higher back transfers compared with the baseline treatment.

Because we are interested in the impact of the implicit reputation cue on the trustees' social preferences, it is important that the trustee chooses his back transfer with the exact knowledge of how much the investor sent. This feature of our design ensures that unknown beliefs about the investors' transfers do not affect the trustees' choices. In this respect, our design differs substantially from the public goods experiments of Bateson et al. (2006) and Burnham and Hare (2007), because it is not clear why subjects change their contributions in response to a cue in a public goods experiment. In principle, the cue could cause a more optimistic belief about the other players' public good contributions, which will then lead to an increase in the subject's own contribution; it is known that many subjects are conditional cooperators (Fischbacher et al., 2001; Kurzban and Houser, 2005; Croson, 2007; Kocher et al., 2008), that is, they are willing to contribute more to the public good if they believe that other group members contribute more. Alternatively, the reputation cue could have a direct impact on subjects' social preferences, implying that subjects are willing to contribute more for any given belief level. If the first hypothesis holds, the reputation cue does not affect subjects'

social preferences; it only renders their beliefs about others more optimistic, which then causes the change in behavior. If the second hypothesis holds, the reputation cue has a direct effect on subjects' social preferences. In our experimental design a change in the trustees' behavior cannot be attributed to changes in their beliefs because the trustees know the exact investment when they make their back transfer. Thus, we can measure the impact of the implicit reputation cue for any given transfer level, which provides a clean behavioral measure of a change in social preferences.

Because we measure subjects' degree of Machiavellianism and their fairness standards, we are able to examine whether subjects who score differently on these measures respond differently in the different treatments. We expect, in particular, that highly Machiavellian subjects tend to transfer back less in the baseline condition. It is also important to examine the impact of the implicit cue condition for subjects who score high and low on the Mach score. In particular, if the non-selfish subjects (i.e., those scoring low on the Mach score) are particularly responsive to the implicit reputation cue, one may be more inclined to attribute the observed prosociality in anonymous one-shot experiments to uncontrolled implicit reputation stimuli. Alternatively, if eye cues trigger reputation concerns, *opportunistic* subjects (i.e., those scoring high in Machiavellianism) could be particularly responsive to such cues. In contrast, if subjects' Mach scores do not affect the response to the implicit cue, one may have more confidence in the hypothesis that the prosocial behavior in anonymous experiments is a true expression of subjects' social preferences and not just an artefact of uncontrolled implicit reputation cues.

In the explicit reputation treatment, the subjects' personal identities are still kept anonymous but we render the history of the trustees' back transfers observable for their current investors. Thus, each investor can assess the past willingness of the current trustee to transfer back resources. Because the trustees know this, even selfish trustees now have an incentive to choose "compensate" or "equalize", because in this way they can increase the likeli-

hood that the investors they face in the future (and know their past choices) will make large investments. This explicit reputation incentive ceases in the final period (when there will be no future encounters with investors): the selfish trustees will defect in the last period, and only the trustees with social preferences will make positive back transfers.

The effectiveness of the explicit reputation incentive requires that the trustees understand that their current back transfers will affect average investments of future investors. Thus, the explicit reputation incentive will only raise the trustees' back transfers if the trustees exhibit this kind of rationality. Reputation incentives can also increase back transfers of subjects with social preferences. They may, for example, choose "equalize" instead of only "compensate" when the pecuniary incentive coincides with their social motive. The hypothesis that explicit reputation incentives increase trustees' transfers is also backed by previous findings (Gächter and Falk, 2002; Cochard et al., 2004).

Our measure of Machiavellianism enables us to examine whether there is a meaningful heterogeneity in trustees' responses to the explicit reputation incentive. In view of the behavioral and neurophysiological evidence documented in Spitzer et al. (2007), it seems plausible to conjecture that highly Machiavellian subjects respond more strongly to the explicit reputation incentive. Future investors are likely to punish low back transfers by lowering their investments. By definition, highly Machiavellian subjects are particularly susceptible to such threats. Therefore, they should respond more strongly to the pecuniary reputation incentives.

2.4 Results

2.4.1 The Implicit Reputation Cue

In this section, we examine the impact of implicit reputation cues on trustees' back transfers. If the implicit cue raises reputation concerns, the trustees in

the implicit cue condition should make higher back transfers than those in the baseline condition. Moreover, if the implicit cue has a sufficiently strong effect, the back transfer pattern in the implicit cue condition should resemble the pattern in the explicit reputation condition. Finally, if the implicit reputation cue raises back transfers, this may also increase investments because higher investments increase the investors' payoffs if a sufficiently high share of trustees choose to equalize payoffs (see the final column in Table 2.1).

Table 2.2: Descriptive statistics of average, median and modal choices across treatments.

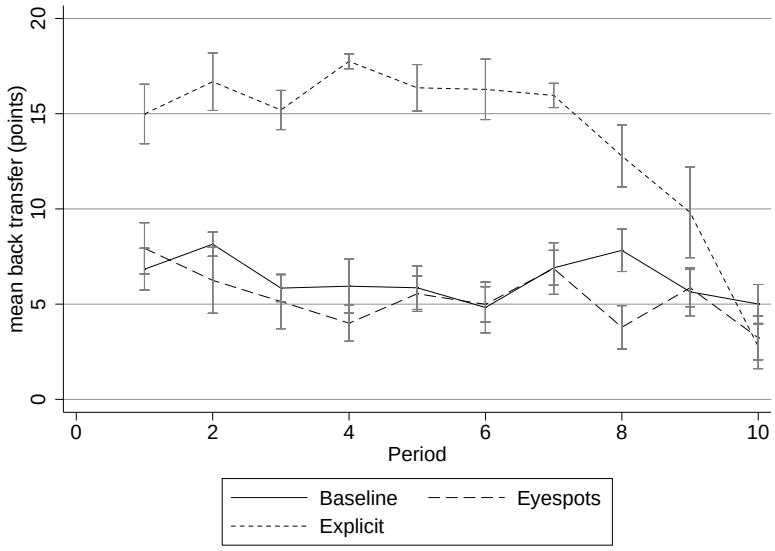
Treatment	baseline	implicit	explicit
average back transfer	6.28	5.36	13.86
median trustee decision	compensate	nothing	equalize
mode trustee decision	nothing	nothing	equalize
average investment	5.88	5.74	7.73
median investment	7	7	10
mode investment	10	10	10
# subjects	108	108	72
# matching groups	9	9	6

Table 2.2 provides a first indication of the impact of the implicit cue condition. In the baseline condition, the average back transfer is 6.28 points and the trustees' modal choice is 'nothing'. The average back transfer in the implicit cue condition is even somewhat lower and the modal choice is also "nothing". The small difference in the means across conditions is not significant ($p = 0.402$, $n = 18$, Mann-Whitney U test). The investors' transfer choices are also very similar across the two conditions. The median investment level in both conditions is 7; investors in the baseline condition invest an average of 5.88, while the average investment in the implicit cue condition is 5.74 ($p = 0.825$, $n = 18$, Mann-Whitney U test).

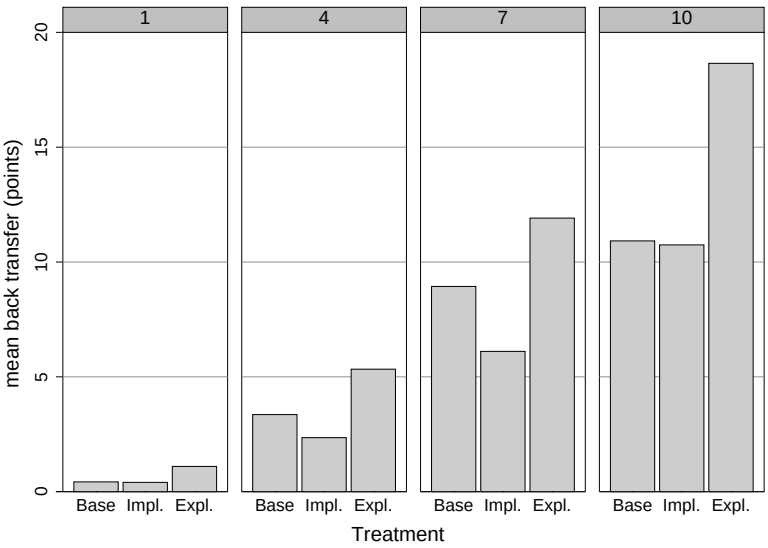
Figure 2.2a shows the time path of average back transfers. The figure indicates that the average back transfer varies between 5 and 8 units in both the baseline condition and the implicit cue condition, with little difference

Figure 2.2: Trustees' mean back transfers.

(a) Over time across treatment conditions; error bars represent standard errors on matching group level ($n = 24$).



(b) Per investment level across treatment conditions.



between the conditions. The figure also displays the standard errors (clustered on matching groups) of the mean, which indicate that the differences between baseline and implicit cue conditions are not significant ($p \geq 0.272$, Mann-Whitney U tests with Holm-Bonferroni correction for multiple comparisons). Thus, Table 2.2 and Figure 2.2a provide little indication that the implicit cue condition increased average back transfers.

An examination of the impact of implicit cues in more detail requires further control for the investments that the trustees face. Figure 2.2b shows the trustees' average back transfer conditional on the received investments. On average, trustees in the implicit cues condition sent back the same or a slightly smaller amount than in the baseline condition for any given investment level. In Table 2.3 we report the results of ordinary least squares regressions with the average relative back transfer as the dependent variable. The dependent variable is relative back transfer, defined as the share of points returned over points received (i.e., the quadrupled investment). Thus, a choice of "nothing" is equivalent a relative back transfer of 0 percent, the choice of "compensate" translates into a relative back transfer of 25 percent and the choice of "equalize" means that 62.5 percent of the received points are sent back by the trustee. For example, with an investment level of 4, the trustee receives $4 \times 4 = 16$ points and sends back 10 if he chooses "equalize", giving a payoff of 16 to each of the two players; the relative back transfer equals $10/16 = 62.5$ percent. Likewise, if the investor sends 10, the trustee receives 40 and sends back 25 in the case of "equalize", which yields a relative back transfer of $25/40 = 62.5$ percent. The advantage of using relative back transfers is that a given choice, such as "equalize", implies the same percentage number regardless of the investment level. Thus, our regressions implicitly estimate the conditional frequency of the three choices "nothing", "compensate" and "equalized". The regressors are described in Table 2.4.

Model (1) in Table 2.3 reports the result of a regression that takes the average relative back transfer per matching group as the dependent variable. The independent variables in this regression are the average investment per

Table 2.3: OLS regression analysis of trustee decisions. Models (1) and (2) use aggregated data on the matching group and trustee level, respectively, i.e., the variables are averages on the matching group and trustee level, respectively. In model (1), 'Mach' and 'Fairness' represent matching group averages.

Dependent variable: relative back transfer	(1)	(2)	(3)
(Mean of) Investment level	0.031***	0.033***	0.018***
	(0.008)	(0.007)	(0.003)
Implicit	−0.013	−0.040	−0.038
	(0.020)	(0.033)	(0.036)
Explicit	0.170***	0.107**	0.203***
	(0.029)	(0.039)	(0.038)
(Mean of) Mach	0.062	−0.086*	−0.081*
	(0.039)	(0.043)	(0.043)
Mach × implicit		0.061	0.057
		(0.055)	(0.056)
Mach × explicit		0.110**	0.105**
		(0.048)	(0.049)
(Mean of) Fairness	−0.108**	−0.112*	−0.109*
	(0.043)	(0.056)	(0.054)
Fairness × implicit		−0.000	−0.006
		(0.062)	(0.062)
Fairness × explicit		0.007	−0.018
		(0.064)	(0.061)
Explicit × last 3			−0.176***
			(0.032)
Period			−0.006***
			(0.002)
Constant	0.049	0.113**	0.230***
	(0.051)	(0.045)	(0.031)
N	24	144	1440

Standard errors in parentheses (clustered on matching groups in models (2) and (3)).
* p<0.10, ** p<0.05, *** p<0.01

Table 2.4: Description of regressors

Investment Level	number of points the investor transfers
Implicit, Explicit	dummies for the respective treatments (omitted category: baseline treatment)
Mach	dummy variable that equals 1 if the participant scored above the median in the Mach-IV inventory
Fairness	dummy variable that equals 1 if the participant's answer to the fairness question was above the median response (i.e., if the subject's fairness standard is below the median)
Explicit × *last 3*	dummy that equals 1 if the observation comes from periods 8, 9 or 10 in the explicit treatment. Its purpose is to capture the end-game effect that occurs when the future benefits from reputation vanish towards the end of the experiment
Period	denotes the experimental period (from 1 to 10)

matching group, dummy variables for the implicit cue and the explicit reputation treatment, the average Mach score of the trustees in the matching group, and the average response to the fairness question (high answer indicates low fairness norm). In all regressions, the omitted category is the baseline dummy, implying that the constant measures the average relative back transfer in the baseline condition, while the dummy for the implicit cue (explicit reputation) condition measures the difference between the baseline condition and the implicit cue (explicit reputation) condition.

Regression (1) is the most conservative because the unit of observation is average behavior in a matching group, giving us 24 observations in total. We find a highly significant positive effect of the investment level, i.e., higher investments generate higher relative back transfers. For our purposes, the most important result of regression (1) is the small and insignificant effect of the dummy for the implicit treatment. The coefficient for this dummy is close to zero, highly insignificant ($p = 0.533$) and even has the "wrong" sign, indicating that eye spots certainly have no positive effect on trustees' back transfers. In addition, we find a significant ($p = 0.021$) effect of the

fairness standard in the baseline condition, i.e., subjects with a lower fairness standard tend to make lower back transfers.

In regression (2), we examine the mean relative back transfer on the individual level. This yields 144 observations, as we have 36 trustees in the explicit reputation treatment ("explicit") and 54 in each of the other two conditions (standard errors are clustered at the matching group level). In models (2) and (3), Mach score and Fairness represent dummies for subjects with an above-median Mach score and a below-median fairness standard; recall that subjects with a below-median fairness standard are those who give a high (above-median) rating on the fairness question. Interestingly, this has little effect on the impact of the investment level and implicit cue condition: we get virtually the same result as in model (1), with respect to both the size and the significance of these coefficients. In particular, the coefficient of the implicit cue treatment is still very small, insignificant ($p = 0.242$) and has the wrong sign. However, because of the larger number of observations, the fairness standard and individuals' Mach score is now almost significant at the five-percent level (both $p = 0.056$); subjects with lower fairness standards and a higher Mach score transfer back less in the baseline condition. We are also able to examine the interaction between the fairness standard, the Mach score and the implicit cue condition in regression (2). The interaction between the fairness standard and the implicit cue condition is clearly insignificant ($p = 0.998$); the same holds for the Mach score ($p = 0.279$). This indicates that the implicit cue condition also does not cause behavioral changes in trustees with different fairness standards and different Mach scores.

Finally, we take the decisions in each period as units of observation and cluster again on matching groups in model (3). The dependent variable is now the individual relative back transfer in a period, which limits the observations to 0, 25 or 62.5 percent of the received amount. In model (3), we also include variables that capture time effects.

The results of model (3) are interesting in several respects. First, and most importantly, the coefficient for the implicit cue treatment remains small

in magnitude and insignificant ($p = 0.304$), and again has the wrong sign. Second, subjects who have a low fairness standard contribute less in the baseline condition ($p = 0.054$). Third, subjects with a high Mach score also contribute less in the baseline condition ($p = 0.076$). Both the second and the third effect are substantial, reducing the mean relative back transfer by between 8 and 11 percent. Fourth, the interaction between the implicit cue condition and the below-median fairness standard/above-median Mach dummy is not significant ($p = 0.921$ and $p = 0.319$, respectively), indicating that individuals with a low fairness standard/high Mach score do not respond differently to the implicit reputation cue compared with individuals with a high fairness standard/low Mach score. Thus, there is no evidence that individuals who score low on selfishness and opportunism are more prone to implicit reputation cues. Both high and low Mach individuals show little response to the implicit reputation cue.

2.4.2 Explicit Reputation Incentives

In this subsection, we examine the effect of pecuniary reputation incentives on the trustees' back transfers and the investors' transfer. Table 2.2 shows that—in contrast to the implicit reputation condition—the explicit reputation condition causes an enormous increase in average back transfers, from 6.28 to 13.86. While the modal response in the baseline condition is "nothing", the modal response in the explicit reputation condition is "equalize". This big change in the trustees' back transfers is highly significant ($p = 0.002$, Mann-Whitney U test) and led to a significant increase in the investors' transfers, from 5.88 to 7.74 ($p = 0.006$, Mann-Whitney U test). In the explicit condition, the maximum investment also represents the median investment choice. This strong impact of pecuniary reputation incentives can also be seen in Figure 2.2a. The average back transfer is much higher in the explicit reputation condition in all but the last few periods. The time path of the average back transfer in Figure 2.2a also indicates the relatively high degree of rationality that seems to be present in our experiment. During the early periods, a

high back transfer generates a good reputation for many remaining periods, implying that the pecuniary return of a good reputation is high. During the final few periods, a high back transfer generates a good reputation only for a few remaining periods, implying that the pecuniary return of a good reputation is lower. Thus, individuals who understand this should choose lower back transfers during the final few periods because the selfish returns of behaving in this way are lower. The time pattern of back transfers in Figure 2.2a is consistent with this rational choice argument.

Interestingly, in period 10 of the explicit reputation condition—in which there are no pecuniary reputation incentives at all—the average level of back transfers is very similar to the level in the other two conditions (in which explicit reputation incentives are absent by design). Thus, the trustees seem to understand the logic of pecuniary reputation incentives quite well: while they do not respond to merely implicit reputation cues that carry no explicit incentive power, they respond strongly to explicit reputation incentives, and they seem to understand when they can gain from a good reputation and when not.

The powerful effect of pecuniary reputation incentives can also be seen in Figure 2.2b: at every investment level, trustees' back transfers are higher than in the other two conditions.

Finally, the regressions in Table 2.3 provide further statistical support for the large effect of the explicit reputation condition. In models (1) and (2), the explicit reputation incentive increases the average relative back transfer by 17.6 and 16.2 percentage points, respectively ($p < 0.001$ and $p = 0.012$). Note that in model (3) the inclusion of the "explicit $\times$ last 3" interaction implies that the "explicit" variable captures the effect of the explicit reputation incentive for the first seven periods while the variable "explicit $\times$ last 3" measures the decrease of back transfers during the final three periods. The coefficient of 0.203 ($p < 0.001$) for the variable "explicit" thus indicates that in the first seven periods subjects increase the relative back transfer relative to the baseline condition by 20.3 percentage points if they face an explicit

reputation incentive. Moreover, highly opportunistic trustees (above-median Mach score) show an increase in relative back transfers that is 10.5 percentage points larger when they face an explicit incentive (coefficient of 0.105; $p = 0.044$). Taken together, these results indicate a large effect of the explicit reputation incentive, an effect that contrasts sharply with the null effect of the implicit reputation cue. In fact, an F-test indicates that the difference between the coefficients of the implicit and the explicit condition is highly significant ($p < 0.001$).

Subjects scoring high on the Machiavellianism scale exhibit lower back transfers in the baseline condition than the low Mach subjects, but they also respond more strongly to the explicit reputation incentive. This response pattern then raises the question whether there is a difference in the back transfer *level* between low Mach and high Mach subjects in the explicit reputation condition. We tested the Null hypothesis of no difference in the level of relative back transfers with an F-test and found that the hypothesis cannot be rejected ($p = 0.312$). Thus, in the presence of explicit reputation incentives, the back transfer behavior of low and high Mach subjects is indistinguishable because the latter compensate for their lower back transfers in the absence of an explicit incentive with a higher response to the explicit incentive.

The above results confirm the hypothesis that Machiavellian subjects respond particularly strongly to social punishment threats such as loss of reputation. This finding is consistent with the results of another study (Simpson and Willer, 2008) that also observes that egoistic subjects show a stronger response to pecuniary reputation incentives.

2.5 Conclusion

There is little disagreement among researchers that explicit reputation incentives strongly affect human prosocial behavior. These explicit incentives can take the form of higher future material benefits in a dynamic experimen-

tal game—such as in our explicit reputation treatment—or they can arise when real people (e.g., an audience) saliently observe other people's cooperative or non-cooperative behavior (Gächter and Fehr, 1999; Rege and Telle, 2004; Kurzban et al., 2007; Smith et al., 2009). The strength of merely implicit reputation cues, in which subjects cannot really acquire a good or bad reputation, is, however, much less investigated.

Therefore, we examined the impact of such cues on the strongly reciprocal behavior of trustees in a trust game. Previous work has argued that eye cues activate reputation concerns, but did not compare the effect of eye cues with the effect of explicit pecuniary reputation incentives. If reputation concerns shaped humans' altruistic inclinations in ancestral environments to the extent suggested in some of the recent literature (Haley and Fessler, 2005; Burnham and Hare, 2007)—that is, if humans are indeed so sensitive to reputation cues that they respond to them even if the cues carry, in fact, no real pecuniary incentive power—subjects should generate patterns in the eye cue condition that resemble the effects of explicit pecuniary reputation incentives.

However, our results indicate that eye cues, which have been hypothesized to represent reliable indicators of potential observability of one's behavior over the course of human evolution, have no effect at all on the trustees' altruistic behavior. The effect of the implicit cues treatment is close to zero, highly insignificant and even has the wrong sign. Moreover, this null effect holds regardless of whether we examine the response of subjects who score high or low on the Mach scale. Our results therefore suggest an extremely cautious view of claims that most of the observed prosocial behavior in anonymous one-shot games should be attributed to uncontrolled implicit reputation cues. At the current state of our knowledge, this claim represents no more than a speculation, lacking empirical support. If it were indeed the case that uncontrolled reputation cues are so important, behavior should also respond to experimentally controlled implicit reputation cues.

The null effect of the implicit reputation cue contrasts sharply with the large impact of explicit pecuniary reputation incentives on trustees' behavior.

The large contrast between the implicit and the explicit reputation condition reinforces our conclusions above. The effect of the implicit cue does not even resemble the effect generated by the pecuniary reputation incentive, suggesting that implicit cues are a relatively weak force.

We also found important individual differences in subjects' responses to the pecuniary reputation incentive. Subjects who score high on the Mach scale behave less altruistically in the baseline treatment, but they respond more strongly to the pecuniary reputation incentive.

Why do other studies find an effect of eye cues on prosocial behavior while we find none? With regard to the studies of Bateson et al. (2006) and Burnham and Hare (2007), the following feature of their experiments might have caused the difference. Both experiments investigate contributions to a public good. As shown above, many people are conditional cooperators, and their contributions therefore depend on their beliefs about other people's contributions. Eye cues could generate more optimistic beliefs about other subjects' cooperation behavior, which then induce higher cooperation rates among subjects with preferences for conditional cooperation. This contrasts with our study in which we have full control over subjects' beliefs because the trustees know exactly the investment level if they make their back transfer. Therefore, in our study, eye cues cannot affect beliefs about other subjects' behavior.

With regard to the study of Bateson et al. (2006)—a field experiment about voluntary contributions to an honesty box in a university coffee room—another feature is also potentially important. Subjects often consume coffee jointly and observe whether their colleagues pay for the coffee. In this case the subject's real reputation—and not just its imagined reputation—is at stake. If eye cues draw attention to the moral appropriateness of paying for one's coffee, then this real reputation incentive may be greatly strengthened. Thus, it is possible that the eye cues in the Bateson et al. (2006) experiment enhanced the already prevailing incentive to maintain one's reputation as an honest coffee consumer. This feature of the Bateson et al. (2006) experi-

ment also contrasts with our experiment because we rule out any interaction between the eye cue and the pecuniary (explicit) reputation incentive.

Why did the eye cue affect the behavior in the dictator game experiment of Haley and Fessler (2005) while lacking effect in our trust game? A possible reason for this may be that the dictator game constitutes a less robust situation. Experimental economists now generally acknowledge that the dictator game is likely to involve more experimenter demand effects (Bardsley, 2008) and is less robust than other games in which subjects interact with each other (Cooper and Kagel, 2010). Therefore, relatively weak forces can affect behavior in the dictator game. Perhaps the implicit reputation cue is one of these weak forces.

> There is no single factor in the whole field of labor relations that does more to break down morale, create individual dissatisfaction, encourage absenteeism, increase labor turnover and hamper production than obviously unjust inequalities in the wage rates paid to different individuals in the same labor group within the same plant.
>
> William H. Davis, National War Labor Board 1942–1946
> in Kochan and Barocci (1985)

Experiment 3

The Relevance of Social Comparison and Strong Reciprocity for Economics

Summary This chapter reports a field experiment to test whether workers respond to wage cuts and whether their response depends on coworkers' wages. Workers were organized in teams of two and paid a flat wage. Either one or both workers in a team suffered a wage cut, while workers in the control treatment continued to earn the initial wage. We show in a difference-in-differences analysis that cutting both workers' wage reduced work performance significantly. However, cutting only one worker's wage resulted in a decrease in performance that was twice as large. In contrast, the spared worker's performance remained unaffected. These findings corroborate the fair wage-effort hypothesis, which can explain involuntary unemployment and wage compression.

3.1 Introduction

Standard economic models of labor markets assume that workers respond exclusively to economic incentives and only care about the absolute level of income. These models do not take into account that fairness motives and social comparison may determine worker behavior. The relevance of these motivational forces, however, was long emphasized in social psychology (Festinger, 1954; Homans, 1961; Adams, 1963) and sociology (Davis, 1959; Runciman, 1966; Pollis, 1968). Textbooks on personnel management—like Kochan and Barocci (1985)—also regard the need for fair and equitable treatment of workers as obvious. These ideas have since been partially integrated into the economic theory of labor markets. A well-known example is the fair wage-effort model by Akerlof and Yellen (1990) which states that workers' fairness concerns constrain firms' wage setting. The model draws on two key assumptions: first, workers withhold effort when they perceive that they are paid less than the fair wage, and second, the fair wage depends on the wages paid to the coworkers in the same firm. Under these assumptions, the firm's wage setting may lead to involuntary unemployment and wage compression.

In this chapter we report evidence from a randomized field experiment exploring whether workers respond reciprocally to wage cuts and the extent to which wages paid to coworkers influence their responses. We conducted the field experiment in collaboration with a firm that sells a card permitting customers to attend parties at selected bars and nightclubs. The firm hired workers for two weeks to sell a promotional card at a price of 5 € or in exchange for customer data. Workers had to work in teams of two.[1] In the pre-intervention week, all workers earned a flat base wage. In the post-intervention week, teams were randomly assigned to one of three treatments: in the control treatment, workers continued to earn the base wage; in the second treatment, both workers in a team suffered a wage cut; in the third treatment, only one randomly chosen worker in a team suffered a wage cut,

[1] Team members worked independently of each other, i.e., we observe the amount of sold cards for each worker individually.

while the other worker continued to earn the base wage. To rule out reputation incentives, workers' employment was limited to two weeks without the prospect of future employment.

We show that workers' performance largely depends on the wages paid to coworkers. When only one worker in a team suffered a pay cut while the coworker was spared, performance of the worker with the lower wage declined sharply by 34 percent. In contrast, when a worker suffered the wage cut along with the coworker, work performance only decreased by 15 percent. This difference in responses is highly significant. When workers were spared from the wage cut while their coworker was not, the spared workers' performance was not affected.

To our knowledge, the present study is the first field experiment to examine the causal effects of wage differentials on worker performance. Empirical evidence on the relevance of social comparison in the workplace has only recently emerged, and the results are mixed. Previous findings stem primarily from laboratory experiments. Aside from the debate whether results from laboratory settings can be generalized to the field (Falk and Heckman, 2009), laboratory experiments entail additional issues when studying social comparison.

For example, natural and salient references for comparison are poorly induced because subjects in laboratory multi-worker firms typically do not interact with each other on a personal level. It is thus not particularly surprising that they often do not compare themselves with their coworkers in this case. For example, in Gächter et al. (2008), subjects did not interact, other than observing anonymous coworkers' wages and effort levels before choosing own effort. As a result, subjects did not respond to coworkers' wages and effort levels. In our field experiment, we created a natural and salient person for comparison by forming teams of two workers who were employed together on two consecutive weekends.

Implementing unequal wages in the laboratory is problematic when subjects play the role of the firm and workers are essentially identical. Wage

differentials are then rarely observed because there is no basis for discrimination. To address this problem, Charness and Kuhn (2007) introduced productivity differentials among workers. Although unequal wages were common, workers did not respond to wages paid to coworkers. Unequal wages might not be perceived as unfair when workers know that they differ in productivity. Instead of implementing productivity differentials, Thöni and Gächter (2009) applied the strategy method to elicit responses to wage differentials. Their results lend some support for social comparison, but are inconclusive on the whole. In our study, we exogenously introduced wage differentials that were not justified by differences between workers. In fact, any justification of wage differentials mitigates social comparison effects. For example, Hennig-Schmidt et al. (forthcoming) justified unequal wages by informing one group of workers that less money was available for them than for the other group. This explanation may have made workers consider wage differences acceptable, therefore precluding a social comparison effect because the fair wage-effort model posits that only wage changes that also affect wage fairness induce effort variations. For this reason, we purposely left wage differentials completely unjustified.

Non-experimental field studies are not better suited for identifying social comparison effects because factors unobservable to the researcher may determine both wage differentials and effort choices (Shearer, 2003).[2] We address this endogeneity problem by randomizing wage differentials.

Much of our previous understanding about the harmful effects of wage cuts on work motivation relies on laboratory experiments and manager interviews. Fehr and Falk (1999), for example, conducted a laboratory experiment to investigate downward wage rigidity in competitive labor markets. They find that firms fear low and unprofitable effort levels as a consequence of low wages and are therefore reluctant to accept low wage offers from previously unemployed workers. As a result, wages settle above the competitive level.

[2]Torgler et al. (2008) investigate the relevance of social comparison among basketball and soccer players. They find that both situations – earning more *and* earning less than teammates – reduce performance.

Bewley (1999) provides an example for an interview study on wage stickiness. He reports that managers are reluctant to cut pay during a recession because they are afraid of the negative effects on work morale. Although interview studies are suggestive, the findings only reflect managers' opinions and not worker behavior.[3]

In a non-experimental study, Lee and Rupp (2007) examine the effects of wage cuts on flight delays in the airline industry. They find that wage cuts trigger hostile responses only when the cut is perceived to be unfair. In the same way as for wage differentials, non-experimental field studies on wage cuts do not solve the endogeneity problem. In addition, non-experimental field studies examine worker behavior in ongoing employment relations, making it difficult to disentangle fairness motives from reputation incentives (Howitt, 2002).[4]

To date, only Kube et al. (2010) provide clean evidence on the negative effects of wage cuts on worker performance. In their field experiment, workers reduced their performance significantly when the wage was lower than the promised wage.

3.2 Experiment

3.2.1 Experimental Procedure

Economic Environment

We conducted the field experiment in collaboration with a German firm that operates a nightlife online portal and sells a card that permits attendance to parties in selected bars and nightclubs. In 2008, the firm initiated a promotion to increase its brand awareness. For this promotion, it hired workers to sell a promotional card in public places and nightclubs.

[3]For example, managers may be reluctant to cut wages because they would have to deal with offended workers who might contest the situation; this would not necessarily imply, however, that the workers would work less.

[4]For example, when firms cut wages, workers may simply punish the firm with lower effort as part of an equilibrium trigger strategy.

Recruitment

Workers were hired over a job market database that listed workers with experience in promotion jobs. They were unaware that they were participating in an experiment. To ensure this, we excluded applicants who knew someone from the firm's permanent staff. Hires received a guarantee that they could work on two consecutive weekends. In order to eliminate reputation incentives, it was made clear that there was no prospect of further employment at the firm.

Upon arriving for the training session, workers were randomly allocated to teams of two and randomly assigned to be either "worker 1" or "worker 2" in a team. Both workers, however, had completely identical tasks and responsibilities. By forming teams of two workers who worked together on two consecutive weekends, we created a natural and salient reference for comparison, namely the coworker in a team.

Workers were then made familiar with the objective of the promotion, its structure and procedures, and with the equipment and clothing. Furthermore, they were trained how to approach potential customers.

Task

Workers' task was to sell the promotional card at a price of 5 € or in exchange for a customer's personal information.[5] Teams were assigned a fixed point-of-sale which was either a shopping avenue or a nightclub. Working hours were Friday and Saturday from 5pm to 8pm for public places and 11pm to 2am for nightclubs.

During these three-hour shifts, workers were mostly on their own and thus had full discretion over the amount of exerted effort. Points-of-sale provided

[5]Customer information was recorded in a database, and invitations to join the online platform were sent to the customers. False information could be identified and attributed to the worker who had acquired it. Workers did not know, however, that the correctness of customer information would be verified.

an attractive opportunity to shirk because workers could unobtrusively converse privately. In case of low sales, they could always claim that nobody wanted the card.

Each team was managed by a team leader, who met the workers before and after work shifts. Team leaders supplied workers with promotional cards, assessed the points-of-sale (for example, number of club visitors), and looked after the workers once or twice per shift. After the shifts, they collected the revenues, customer information, and the remaining cards.[6]

3.2.2 Experimental Design

Treatments

We implemented a differences-in-differences setup with a pre- and a post-intervention week and four treatment groups (HH, LL, HL1, and HL2). The pre-intervention week permits the measurement of workers' baseline performance, thus controlling for worker heterogeneity. This baseline is important because ability and therefore performance may vary strongly across workers.

In the pre-intervention week, all workers earned an hourly base wage of 12 €. For the post-intervention week, we randomly assigned teams to one of three treatments. In the control treatment, "HH", both worker 1 and worker 2 continued to earn the hourly base wage of 12 €.[7] In the second treatment, "LL", both worker 1 and worker 2 suffered a wage cut down to 9 € per hour. In the third treatment, "HL", only worker 2 suffered a wage cut down to 9 € per hour, while worker 1 still earned the base wage of 12 € per hour (see Table 3.1 for a summary of the treatments). Thus, treatment HL consists of

[6] Team leaders were permanent employees of the firm and each of them was responsible for two to three teams. They received a comprehensive set of instructions about their communication with the workers and the handling of potentially problematic situations. In particular, they were instructed to treat all workers in the same manner and were prohibited to motivate or rebuke individual workers. Excerpts from worker and team leader instructions are reproduced in appendix C.1.

[7] Team leaders used the phrase "You continue to earn 12 € per hour. This was the manager's decision." and analogous phrases in the other treatments, see appendix C.1.

Table 3.1: Hourly wages (in €)

Treatment	HH		LL		HL	
Worker	1	2	1	2	1	2
Pre-intervention	12	12	12	12	12	12
Post-intervention	12	12	9	9	12	9

two groups: group "HL1" was composed of workers 1 in treatment HL, while "HL2" was made up of workers 2.

Team leaders did not give a reason for the wage cut to prevent altering the wage that workers consider fair. For example, workers' notion of a fair wage could have changed if team leaders had told that the firm faced the risk of bankruptcy. In addition, devising a cover story would have deceived the workers, which would have been a departure from standard convention among experimental economists.

Design Specifics

Three important aspects of this field experiment need to be stressed. First, we implemented a wage cut so that workers earned, on average, at least the promised wage. Thus, we initially raised all workers' hourly wage from 10 € to 12 €.[8] This avoids ethical concerns associated with experimental pay cuts. In addition, the wage increase helps prevent an attrition bias. Drop-outs after a wage cut would be uninformative because they could be interpreted either as a hostile response or the choice of an outside option because the wage had fallen below a worker's reservation wage. This initial wage increase, however, is associated with the potential cost of mitigating the treatment effects.

Second, we adapted the organizational structure of the promotion to maximize the number of subjects. We ran the promotion twice, hiring different

[8]Upon being hired, workers were promised an hourly wage of 10 €. Then, at the beginning of the first shift, team leaders informed the workers about the wage increase. If workers asked for a reason, they were told that the manager had made the decision.

workforces each time.[9] In addition, we ran the promotions in two cities and at two different types of points-of-sale in each of the promotion drives. This made some balancing constraints on the treatment assignment desirable.[10] We assigned treatments evenly to the two promotion drives, cities, and types of points-of-sale in order to minimize time, city, and location type-specific differences across treatments. To maximize statistical power, each of the four treatment groups comprised the same number of workers (i.e., treatment HL comprised as many workers as treatments HH and LL together). We also ran treatment HL at each point-of-sale: the first time at half of the points-of-sale, and the second time at the other half (i.e., each point-of-sale was assigned to treatment HL once and to either treatment HH or LL the other time). Furthermore, we stratified treatment assignment by gender, and also formed same gender teams to avoid confounds.[11]

Third, we allocated workers who knew each other to the same treatment, thus preventing treatment contamination, i.e. communication among workers from different treatments. However, we separated friends into different teams to preclude friendship arrangements within teams. In addition, allocation of teams in space and time was arranged so that teams from different treatments could not possibly meet.

3.3 Hypotheses

Using a simple framework, we analyze how workers respond to wage cuts and how their response depends on the wages paid to their coworkers. Consider a firm that employs two identical workers—worker 1 and worker 2—for a one-time job and pays them a flat wage. In return, each worker generates revenue for the firm by exerting costly effort. The firm's payoff per worker is

[9] Workers from the first promotion drive never had contact with workers from the second drive.

[10] Assignment to worker 1 and worker 2 was randomized unconditionally.

[11] For example, HL2 workers in mixed-gender teams might reduce their performance because they felt they were victims of sexual discrimination.

revenue generated by the worker minus her wage. The worker's payoff is the wage minus her effort cost.

Assume first that workers exclusively maximize their own material interest. The prediction is then straightforward: since workers receive a guaranteed wage that is not contingent on their performance, their effort will not respond to a change in the flat wage as long as the wage remains above their reservation wage.[12]

Now assume that in addition to their own material interest, workers care about fairness. Their fairness perceptions can be based on either the actions or outcomes of others.[13] For example, workers could view a wage cut as a hostile act by the firm and as a consequence reduce their effort. If only one worker's wage is cut, he or she could consider this act even more hostile and further amplify the effort reduction. Alternatively, inequalities in outcome might influence worker behavior. Outcome-oriented fairness models have the advantage of being tractable. We therefore use the model of inequity aversion by Fehr and Schmidt (1999) to generate our hypotheses,[14] and to derive the formal predictions in appendix C.2.

Suppose that workers dislike inequity when comparing their own payoff, that of their coworkers, and the profit the firm earns. Not only effort cost, but payoff comparison as well, determines a worker's effort. Higher effort decreases own payoff and increases the firm's payoff, but leaves the coworker's payoff unaffected.

[12] Alternatively, a high wage could be regarded as a disciplining device for selfish workers (Shapiro and Stiglitz, 1984): workers provide high effort to avoid being fired and losing the high wage. After a wage cut, workers decrease their effort because they have less to lose. According to this model, effort decreases towards the end of the employment (end game effect) and a wage cut decreases effort *regardless* of the coworker's wage. This theory, however, is improbable in our setup because there was no threat of firing.

[13] Action-oriented fairness models include Rabin (1993), Dufwenberg and Kirchsteiger (2004), and Falk and Fischbacher (2006); outcome-oriented fairness models include Fehr and Schmidt (1999) and Bolton and Ockenfels (2000); Levine (1998) presents a type-based fairness model.

[14] The purpose of our experiment is to provide causal evidence on the effect of social comparison on work performance and not to discriminate between closely related fairness models.

Because both team members are paid the same wage in treatments HH and LL, they will both provide the same effort to avoid inequity between them. The firm, on the other hand, creates inequity to the workers' advantage by paying a wage to the worker. The higher this wage, the greater the firm-worker inequity. Thus, inequity averse team members will provide the same amount of effort, which will be higher in treatment HH than in treatment LL.

In treatment HL, the firm pays worker 2 a lower wage than worker 1. In this case, payoff comparisons both with the firm and with the coworker affect effort choices. Consider worker 1 first: provided both workers exert the same level of effort as in HH, worker 1 receives the same payoff as the firm but a *higher* payoff than worker 2. Consequently, worker 1 could increase her effort in order to reduce advantageous inequity with respect to her coworker. This, however, would increase not only effort cost but also overall inequity: while an increase in effort decreases the inequality between workers as worker 1's payoff is reduced, it creates a larger inequality between her and the firm through both a reduction in her own payoff and an increase of the firm's payoff. Worker 1 will therefore not exert more effort; as a result, HL1 effort will be the same as in treatment HH.

Consider now worker 2: provided that both workers exert the same level of effort as in HH, worker 2 receives a *lower* payoff than the firm and worker 1. Hence, reducing effort decreases not only effort cost but also inequity with respect to both the firm and the coworker. In order to equalize payoffs with respect to the firm, worker 2 would provide the same low effort level as in LL. Yet, at this effort level, worker 2 still gets a smaller payoff than worker 1.[15] If worker 2 further decreases her effort, she not only saves effort cost but also reduces disadvantageous inequity with respect to her coworker; this comes at the cost of a disparity between her and the firm to her advantage. This cost, however, is small because envy looms larger than compassion. Worker

[15]Recall that a reduction in effort decreases inequity with respect to the firm more effectively than inequity with respect to the coworker because own effort affects the firm's payoff but it does not affect coworker's payoff.

2 therefore provides less effort than when both workers earn the low wage; consequently, HL2 effort is lower than in treatment LL.

In summary, the model generates the following hypotheses about the change in effort from pre- to post-intervention week:

If workers are sufficiently inequity averse, then

(H1) workers in treatment **LL** decrease their effort after the wage cut.

(H2) workers in treatment group **HL2** decrease their effort after the wage cut more than those in treatment LL.

(H3) workers in treatment group **HL1** do not change their effort after the wage cut.

3.4 Results

3.4.1 Descriptive Statistics

Our sample consists of 96 workers in 48 teams. Table 3.2 shows that workers were predominantly women (77 percent) and, on average, in their early twenties (mean age: 20.7 years). All but one were German citizens; 29 workers, however, had a second nationality (mostly Eastern European). Of the 96 workers, three workers got sick before any wage cut was announced and missed out on the entire post-intervention week.[16] No worker, however, dropped out after the wage cut.

In total, workers sold 8750 promotional cards; mean sales were 22.8 cards per three-hour shift and worker. Only 187 customers (2.1 percent) chose to pay 5€ for the card, while the remaining sales were generated by collecting customer information. Of the 8563 sets of customer information, only 191 (2.2 percent) were false.

[16]These workers were replaced by spare workers who were treated in exactly the same way as the replaced workers would have been treated. We exclude spare workers from the analysis, however.

Table 3.2: Descriptive statistics

Treatment Group	HH	LL	HL1	HL2	Total
# Workers	24	24	24	24	96
# Female	18	18	19	19	74
Age (mean)	20.5	21.2	20.2	21.1	20.7
# Reported sick	1	1	0	1	3
Performance (mean)					
Pre-intervention	20.8	22.4	24.3	22.0	22.4
Post-intervention	22.8	21.4	26.5	18.4	22.3
Overall					22.3

3.4.2 Control Variables

We balanced the treatments over variables known in advance (promotion drive, city, point-of-sale, and gender). Two particular factors, however, were impossible to anticipate, namely how many customers the workers would meet at the point-of-sale (demand), and heterogeneity in worker characteristics. Both factors influence worker performance and can cause spurious correlations in the data if they are not accounted for.

Based on their visits during work shifts, team leaders assessed demand on a 5-point scale (-2 = low, 2 = high). We use this assessment to test for systematic variation in demand across treatments. We cannot reject the null hypothesis that demand was equally distributed across treatments ($p = 0.23$, Kruskal-Wallis test).[17] Nevertheless, there is considerable variation in demand (standard deviation: 1.19). In order to estimate the treatment effects more precisely, we include demand as an additional control in the regression analysis.

To control for worker heterogeneity, we implemented a difference-in-differences design which allows us to include individual fixed effects in the regression analysis.[18] Many field studies report substantial heterogeneity in

[17] All p-values in this chapter are two-sided.

[18] As workers were always assigned to the same point-of-sale, individual fixed effects also capture location-specific differences.

worker ability and point out the importance of controlling for this heterogeneity whenever possible (Shearer, 2004; Fehr and Götte, 2007; Mas and Moretti, 2009).

3.4.3 Treatment Effects

Workers could exert effort in two performance dimensions: quantity (cards sold) and quality (correctness of customer information). Our measure of effort is quality-adjusted performance, defined as total number of cards sold minus sales due to customer data that were verified as incorrect.[19] With random treatment assignment, we can estimate the average causal effect of an intervention by comparing pre- and post-intervention differences in performance across treatments. We first conservatively analyze the impact of the treatments on performance by applying non-parametric tests. For convenience, we present quantities as percentages of the average pre-intervention performance $\overline{y}_{pre}$, resulting in the following hypotheses:

(H1) Workers in treatment LL reduce performance compared to those in the control treatment:

$$\frac{y^{LL}_{post} - y^{LL}_{pre}}{\overline{y}_{pre}} < \frac{y^{HH}_{post} - y^{HH}_{pre}}{\overline{y}_{pre}}$$

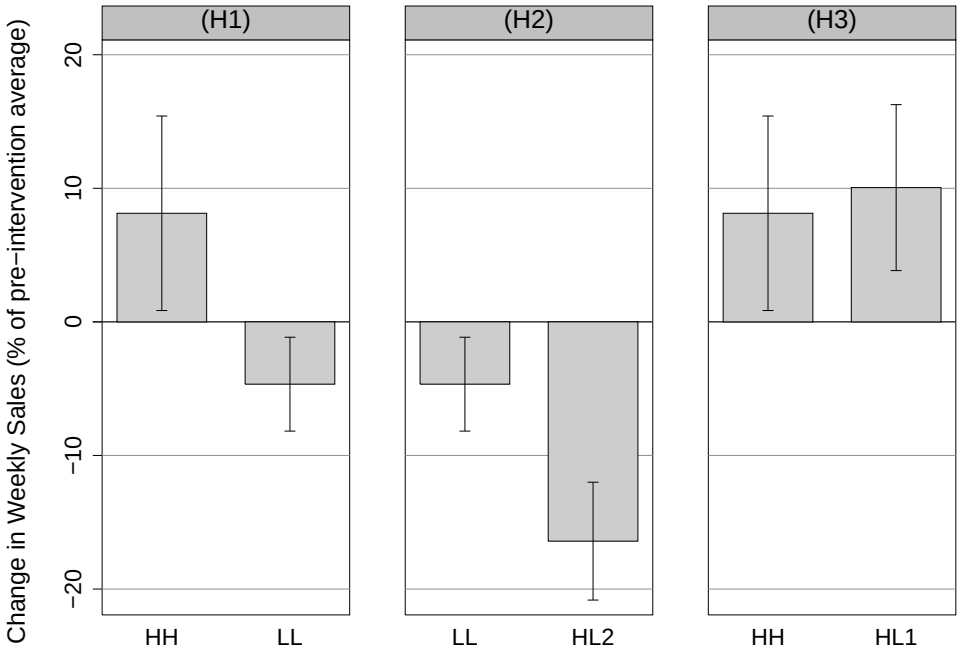
(H2) Workers in treatment group HL2 reduce performance more than those in treatment LL:

$$\frac{y^{HL2}_{post} - y^{HL2}_{pre}}{\overline{y}_{pre}} < \frac{y^{LL}_{post} - y^{LL}_{pre}}{\overline{y}_{pre}}$$

(H3) Workers in treatment group HL1 provide the same performance as

[19] The empirical results do not change if we include incorrect customer data in the analysis.

Figure 3.1: Percentage change in sales from pre- to post-intervention week.
H1: change in treatment groups HH and LL;
H2: change in treatment groups LL and HL2;
H3: change in treatment groups HH and HL1.



(error bars represent standard error, clustered over teams; spare workers excluded)

those in the control treatment:

$$\frac{y_{post}^{HL1} - y_{pre}^{HL1}}{\overline{y}_{pre}} = \frac{y_{post}^{HH} - y_{pre}^{HH}}{\overline{y}_{pre}}$$

Figure 3.1 shows the percentage change in performance from pre- to post-intervention week by treatment. Performance in the control treatment HH increased non-significantly by 8 percent relative to the pre-intervention average ($p = 0.58$, Wilcoxon Signed Rank test), which points to learning effects. Persuant to our hypotheses, we test for differences between this baseline change and changes in the other treatments. Compared to the control treatment, performance in treatment LL decreased non-significantly by 13

percentage points ($p = 0.37$, Mann-Whitney U test), lending weak support for hypothesis (H1). By contrast, a wage cut for only one worker had a dramatic impact on performance: workers in the HL2 group significantly decreased their performance by 24 percentage points compared to the control treatment ($p < 0.01$, Mann-Whitney U test). This reduction in performance is also stronger than the decrease in the LL group ($p < 0.05$, Mann-Whitney U test), providing strong first evidence for hypothesis (H2): workers reacted more drastically to wage cuts when the coworker was spared. Moreover, if we compare the paired observations within HL teams, we see that HL2 workers also reduced their performance compared to their coworkers ($p < 0.01$, Wilcoxon Signed Rank test). Finally, consistent with hypothesis (H3), the spared coworkers hardly responded to the wage cut: HL1 workers increased their performance by 2 percentage points compared to the control group ($p = 0.37$, Mann-Whitney U test).

Our non-parametric analysis does not control for differences in demand and worker characteristics. To address this issue, we estimate a difference-in-differences regression model that uses the balanced panel data structure with each worker as a panel unit and each team as an independent observation:[20]

$$\log(y_{ikt}) = \alpha + \nu_i + \theta_t + \delta D_{kt} + \beta_1 I_{kt}^{LL} + \beta_2 I_{kt}^{HL1} + \beta_3 I_{kt}^{HL2} + \epsilon_{ikt} \qquad (3.1)$$

where $\log(y_{ikt})$ denotes the logarithm of average performance of worker i in team k and week t; the constant α captures the average pre-intervention performance, ν_i represents individual fixed effects, θ_t captures the baseline trend from pre- to post-intervention week, and D_{kt} controls for differences in demand; I_{kt}^{g} are intervention dummies for whether the respective intervention has affected treatment group g in week t (the omitted category is the control group HH); finally, ϵ_{ikt} is an idiosyncratic error term, which is clustered over teams. Recall that individual fixed effects not only capture time invariant differences across workers but also location-specific factors because workers

[20]For an exposition of difference-in-differences estimation, see Bertrand et al. (2004).

Table 3.3: Treatment effects on performance

Dependent variable: log(performance)	(1)	(2)
Post-intervention	0.058	0.090
	(0.080)	(0.074)
LL × Post-intervention	−0.106	−0.145*
	(0.090))	(0.079)
HL1× Post-intervention	0.034	0.015
	(0.095)	(0.089)
HL2× Post-intervention	−0.306***	−0.342***
	(0.103)	(0.101)
Demand		0.117***
		(0.034)
Constant	3.057***	3.048***
	(0.015)	(0.015)
Individual Fixed Effects	Yes	Yes
N	189	179
Adj. R^2	0.202	0.312

Standard errors in parentheses, clustered over teams.
* p<0.10, ** p<0.05, *** p<0.01

were always assigned to the same point-of-sale.[21]

Table 3.3 presents the estimates for the treatment effects. Column (1) presents the results for equation (3.1) omitting the control variable for demand. The "Post-intervention" dummy represents the percentage change in baseline performance from pre- to post-intervention week. The coefficient of this dummy shows that performance in treatment HH increased non-significantly by 6 percent ($p = 0.47$, t-test). The intervention dummies, i.e. the three interactions "$g\times$ Post-intervention", describe how the change in performance differed with respect to the control group. In treatment LL, the change in performance was 11 percentage points lower than in treatment HH; this negative response is twice as large in magnitude as the baseline

[21] As expected, the fixed effects parameters are highly significant ($p < 0.001$, F-test).

change, reaffirming hypothesis (H1). However, data are too noisy to reject the null hypothesis of no difference between treatment HH and LL ($p = 0.25$, t-test). Workers in the HL2 group, however, reduced their performance by 31 percentage points ($p < 0.01$, t-test) compared to treatment HH. They responded three times more strongly to the wage cut than workers in treatment LL ($p = 0.01$, Wald test), which strongly corroborates hypothesis (H2). In line with hypothesis (H3), workers in group HL1 did not significantly increase their performance compared to the control group ($p = 0.72$, t-test).

In Column (2), we include demand as a control variable in order to reduce residual variance.[22] The coefficient of demand is highly significant ($p < 0.01$, t-test) and has the expected sign: the more potential customers, the likelier workers could sell a card. The inclusion of the demand variable does not qualitatively change the results. It allows us, however, to estimate the treatment coefficients more precisely. As a consequence, the influence of the wage cut affecting both workers is now significant at the 10 percent level. The point estimates for the two wage cut interventions imply a performance reduction of 15 percentage points for the LL group ($p = 0.07$, t-test) and 34 percentage points for the HL2 group ($p < 0.01$, t-test). The difference between the LL and the HL2 group is again highly significant ($p = 0.01$, Wald test). The coefficient for the HL1 group remains small and insignificant ($p = 0.87$, t-test).

Taken together, the results in Table 3.3 weakly support hypothesis (H1), but they strongly corroborate hypothesis (H2). The results demonstrate that workers are not only concerned about their own wage. Rather, workers compare themselves with their coworkers and respond more negatively to wage cuts if they end up earning less than their coworkers. Finally, consistent with hypothesis (H3), preferentially treated workers do not increase their effort, pointing out the boundaries of social comparison effects.

[22] When demand is included in the regression, the number of observations decreases by 10 because the demand measure is missing for some shifts.

3.5 Conclusion

This chapter reports evidence from a randomized field experiment investigating fairness motives and social comparison in a real-life employment situation. When both workers in a team suffered a wage cut of 25 percent, performance declined by 15 percent. We may even underestimate this effect. Workers may have perceived the wage cut as less unfair because they initially received a wage increase that partially outweighed the wage cut.

This result provides causal evidence on why firms often refuse to cut wages even though excess labor supply exists and labor markets have not yet cleared. The question about downward wage stickiness has played a key role in a long-lasting macroeconomic debate starting with Keynes in the 1930s. Bewley (1999) was able to show that managers regard fairness as the most important reason for downward wage rigidity. In Bewley's interviews, managers revealed that they were afraid of workers' resentment in response to wage cuts. To date, behavioral evidence for the fairness explanation is still scarce because exogenous wage cuts are rarely observed.[23] We fill this gap by implementing randomized wage cuts in a real-life job of limited duration to exclude explanations other than fairness.

Our main result demonstrates the key role of social comparison in the workplace. When only one worker in a team suffered a wage cut of 25 percent, the affected worker's performance declined, on average, by 34 percent. This effect is much stronger than the effect of the pay cut for all workers, even though the wage was reduced by the same amount. This result provides clear evidence for the conjecture that workers respond to wages paid to their coworkers. In contrast, spared workers did not respond to the wage cut their coworkers suffered. This asymmetric effect speaks to the results of a field experiment by Cohn et al. (2009). They show that workers who feel overpaid do not respond to wage increases, while workers who feel underpaid respond with a performance increase.

[23]The only exception known to us is Kube et al. (2008).

This chapter sheds light on why firms usually avoid paying some workers less than others on the same job. In the 1980s, for example, airline companies such as American Airlines, Delta, and Northwest introduced two-tier wage systems where new hires were paid less than incumbent workers. These wage policies, however, were phased out in the 1990s due to the resentment of the workers as well as the high turnover these lower wages generated (Card, 1997). The New York Times noted that two-tier wage systems have "produced a resentful class of workers who in some cases are taking their hostility out on customers" (Salpukas, 1987). Social comparison may also affect strategic decisions of companies, such as mergers and acquisitions. For instance, the acquisition of Piedmont Aviation by U.S. Airways entailed unexpectedly high acquisition cost due to wage increases at Piedmont to U.S. Airways' more generous salary levels (Kole and Lehn, 2000). Such costly measures are necessary to ensure internal pay equity. When General Electric (GE) acquired NBC in 1986, for example, engineers at GE were angered by the fact that their colleagues at NBC earned higher salaries (Camerer and Malmendier, 2005). Our study provides causal evidence that wage disparities within firms also greatly damage workers' productivity. This provides a plausible reason why firms frequently prefer compensation practices that maintain firm-internal equity such as wage compression (Akerlof and Yellen, 1990) and wage secrecy (Lawler, 1990).

Together, these two findings correspond to the behavioral relation between wages and effort levels described in the phenomenological model by Akerlof and Yellen (1990).[24] It is important to note that other efficiency wage models, such as the well-known shirking model of Shapiro and Stiglitz (1984), cannot explain these results. In particular, this alternative model does not predict that responses to wage cuts depend on coworker wages.

The present experiment focuses on one specific determinant of workers' fair wage, namely coworkers' wages. Although this may be the most important determinant, there may be others. For example, past wages may

[24] As shown in appendix C.2, this behavioral pattern can be derived from recently developed models of social preferences.

influence what workers think they are entitled to (Kahneman et al., 1986). Thus, fairness considerations may also have severe implications for the optimal wage policy over time.

Another important aspect is the communication of wage policies. For example, workers may be willing to accept wage cuts when they feel that they are justified (Greenberg, 1990). Therefore, managers may prevent adverse consequences following from pay cuts if they can thoroughly and sensitively explain the reason for the wage reduction, for example in order to avoid bankruptcy.

Remember upon the conduct of each depends the fate of all.

---

attributed to Alexander the Great

Conclusion

Positive and negative reciprocity shape our behavior in many important ways, Adam Smith writes in his "Theory of Moral Sentiments":

> Gratitude and resentment, therefore, are the sentiments which most immediately and directly prompt to reward and to punish. To us, therefore, he must appear to deserve reward, who appears to be the proper and approved object of gratitude; and he to deserve punishment, who appears to be that of resentment. (Smith, 2010)

Throughout the different chapters of my dissertation, reciprocity came in various guises, as positive or negative reciprocity, as strategic or strong reciprocity. One major driver of reciprocal behavior is the concern for one's good reputation: when future interaction partners have access to information about past behavior, people become more cooperative and accentuate their positive reciprocity.

Experiment 1 shows that the ability to translate reputation concerns into actions is crucially tied to a particular brain region, the right dorsolateral prefrontal cortex. But even in the absence of explicit reputation concerns, people show a substantial level of reciprocity, so called strong reciprocity. Recent studies suggested that even this kind of reciprocity is mainly driven by reputation concerns, albeit only implicitly. Experiment 2 shows that strong reciprocity is not connected to the presence of implicit reputation cues. This result calls certain theories into question that describe prosocial

behavior in anonymous one-shot interactions as a non-adaptation of mental programs of reputation building to modern social environments (Haley and Fessler, 2005; Bateson et al., 2006). The reverse of cooperative behavior and positive reciprocity is punishment out of negative strong reciprocity. Experiment 3 shows that workers display negative reciprocity when they feel treated unfairly. More importantly, the major indicator of wage fairness turns out to be not absolute wage but wage *relative to coworkers*.

This last result has major welfare implications. What people consider as fair determines their actions and produces reciprocity that can enhance efficiency of economic interactions when complete contracting is infeasible. On the other hand, reciprocity can impair social welfare when fairness perceptions prevent firms from setting wages more flexibly and from being transparent about their wage policy. In the worst case, people who would like to work remain unemployed because wages do not adjust to the market clearing level.[25] Economic research must design solutions to reduce involuntary unemployment, and for this purpose, economists have to shed light on the mechanisms that produce it. Experiment 3 highlights the role that field experiments can play in this pursuit.

In order to explore the mechanisms behind reciprocal behavior, I have applied three different empirical approaches—"standard" behavioral laboratory experiments, neuroeconomic laboratory experiments and controlled field experiments—to explore why and how people cooperate. Only the combination of different methods can provide a full picture of the mechanisms of human social interactions.

The Role of Neuroeconomics

The first study demonstrates how neuroeconomic methods can add a new dimension to findings from the standard behavioral laboratory by highlighting the neural underpinnings of observed behavior. Recently, economists have

[25] At the time of writing, 23 million people in the European Union are unemployed (Allen, 2010).

doubted the relevance of behavioral economics and neuroeconomics for empirical economics and economic modeling (Gul and Pesendorfer, 2008, 2007). These doubts rely on the view of economics as a positive science. The normative dimension—welfare economics—is only seen as a vehicle to spark new descriptive research (for example, the finding that an institution is Pareto-inefficient leads to research how inefficient institutions can persist). The notion of welfare as based on the axiom of revealed preferences, according to this line of argument, should not be taken too literally, and economists should abstain from normative policy prescriptions:

> Greater psychological realism is not an appropriate modeling criterion for economics, and therapeutic social activism is not its goal. Welfare analysis helps economists understand how things are by comparing the existing situation to how things might have been in a plausible alternative institutional setting; welfare theory is not a blueprint for a social movement. (Gul and Pesendorfer, 2008, p. 36)

Their bottom line is that "the best way to understand welfare economics is to view it as a part of positive economics" (Gul and Pesendorfer, 2007, p. 475).[26]

While welfare economics in the sense of Gul and Pesendorfer may often be fruitful for finding reasons why inefficient institutions can persist, many inefficiencies may arise precisely because humans are not *homines oeconomici*, because they are biased, use flawed heuristics in reasoning and have preferences that are more intricate than some economists may think. Economists interested in the causes why some institutions are not well-behaved (in the sense of standard economics) therefore need behavioral models of human interactions that depart from the view of revealed preferences as an axiom and develop a richer theory of decision-making.[27] They are also likely to need

[26] A minority of economists actually sees economics based on the axiom of revealed preferences as a normative science, and the profession of economists as a therapeutic one (Landsburg, 2007; Friedman, 1995). They advocate libertarian ideologies and policy recommendations.

[27] Bernheim and Rangel (2007) provide an exposition of "behavioral public economics".

neuroeconomic methods because these methods could provide a tool for deciding in which situations a theorist can safely make the assumption that choices reveal preferences and when this assumption must be rejected—in other words, when it is safe to equate "decision utility" with "experienced utility" and when it is not.[28] And in cases where the revealed preferences assumption has to be rejected, neuroeconomics can also help distinguish between different competing behavioral theories.

Besides being of scientific interest in their own right, insights into the biological principles of human decision-making can therefore inform normative judgment by bridging the gap between preferences and choices. By removing the need to infer preferences from choices, they can lead the way to a true therapeutic purpose for economics. In other words, if neuroeconomic studies reveal that a certain behavior is not the rational consequence of a preference, but that a neurophysiological mechanism has caused the action to be inconsistent with the preference, then therapeutic or paternalistic measures to these persons' own good may be justified.[29]

In chapter 1, for example, I have described how participants were physiologically manipulated to act against their own better judgment. This finding should make researchers more cautious to flatly equate "decision utility" with "experienced utility"—although this abstraction is often useful—and to derive premature policy recommendations from what seem to be revealed preferences. Another example comes from my ongoing research (data unpublished). It seems to be the case that the activity of a certain brain region—the amygdala—is causally involved in the amount of trust one displays towards another person. We show that subjects whose amygdala activity is exogenously increased behave less trusting than control subjects. One could thus speculate that people who display a very high amount of trust towards others due to a hypofunctional amygdala should be protected from their naïveté because their behavior does not reveal a preference for trusting others. Rather,

[28] See Plassmann et al. (2007) for an example.

[29] An attempt to reconcile freedom of choice with therapy is minimally-coercive or "libertarian" paternalism (Thaler and Sunstein, 2003; Loewenstein and Haisley, 2008).

it is just a symptom of a malfunctioning decision mechanism. In short, neuroeconomics is not a rival of standard economics; it is a complement that promises to add a true normative component to economic thought.

The Role of Lab and Field Experiments

The last chapter of this dissertation applies the experimental approach to a field setting. Recently, a debate about the virtues and vices of laboratory and field experiments has been spawned. The economic laboratory has been criticized for using unrepresentative and self-selected subjects, for using artificial environments and being prone to the Hawthorne Effect while field experimentation has been portrayed as the remedy to these flaws (see for example Levitt and List (2007)). Falk and Heckman (2009) reply that while some of these charges may apply to lab experiments, first, they can be largely addressed through well-designed control treatments, and second, they apply to field experiments to much the same extent as to lab experiments. They point out that in field experiments subjects may not be more representative than in the lab, and that environments in field experiments are as specific as those in laboratory experiments. They conclude that different "empirical methods and data sources are complements, not substitutes" (Falk and Heckman, 2009, p. 537). One example for complementarities is that "the field offers a large range of variations in X [environmental parameters, FS], which are potentially relevant but hard to implement in the lab" while the lab offers tight control (*ibid.*).

It is the purpose of the field experiment reported in chapter 3 to provide evidence for social comparison and strong reciprocity in an ecologically valid environment. In an experiment, it is crucial to create the relevant conditions because environmental parameters may influence the existence and size of a behavioral effect of a treatment. As explained in chapter 3, however, it is hard to create the environment needed to observe this effect in the lab. In contrast, the "natural" employer-employee relationship that we examine in our field experiment is ideally suited to test Akerlof and Yellen's (1990)

fair wage-effort hypothesis. Experiment 3 is—to date—the only unequivocal evidence for the existence of a social comparison effect. However, the study is not meant to disentangle different candidate models of social comparison. The discrimination between similar models requires exact measurement of behavioral parameters under tightly controlled conditions, in other words, a laboratory experiment.

By showing that people outside the laboratory strongly reciprocate both positively and negatively in environments that are economically important to them, chapter 3 attests the economic relevance of social preferences. Moreover, it reveals that many situational parameters—even if they are irrelevant for own payoff, like the payoff of another person—may influence how people assess the fairness of the situation and how they will react. This makes building accurate economic models difficult. Growing field evidence like experiment 3, however, suggests that traditional economic modeling will remain imperfect in predicting micro- and macroeconomic outcomes until it augments the stark conception of *homo oeconomicus* to accommodate the findings of behavioral and neuroeconomics that are able to capture the complex motives of social behavior. The *homo oeconomicus* model of decision-making is oftentimes a too simplistic view of the "conduct of each", and thus, traditional economics may end up predicting the wrong fate of all.

Appendices

Appendix A

Appendix to Experiment 1

A.1 Specifications of rTMS Implementation

rTMS was administered to the DLPFC for 15 min before subjects participated in the trust game ("off-line paradigm") (Robertson et al., 2003), using a Magstim Rapid Magnetic Stimulator and a commercially available figure-of-eight coil (70-mm-diameter double-circle, air-cooled). For stimulation of the right and left DLPFC, the TMS coil was placed over F4 and F3 using the electroencephalogram 10-20 coordination system, as in previous studies (Wout et al., 2005; Koch et al., 2005; Grišková et al., 2007). We chose this approach because no previous fMRI data on this paradigm exist. We used the real-time neuronavigation option for BrainVoyager QX 1.6 with the Zebris CMS20S measuring system for real-time motion analysis (Zebris Medical GmbH) to ensure correct placement of the TMS coil during the stimulation. The stimulation intensity was set at 54% of the maximum stimulator output. The coil was held tangential to the subject's head with the handle pointing rostrally. Subjects received a single 15-min, 1-Hz rTMS train (900 pulses) over either the left DLPFC or right DLPFC, or sham stimulation using a Magstim placebo coil, which looks identical to the real coil and also delivers the characteristic "click" sound. Half of the subjects in the sham stimulation group received sham rTMS over the right DLPFC, and half received it over

the left DLPFC.

The rTMS parameters are well within currently recommended guidelines (Wassermann, 1998) and result in suppression of excitability of the targeted cortical region for several minutes after completion of the rTMS train (Robertson et al., 2003). Subjects performed the task immediately after the end of the stimulation train in the same laboratory room. Because the subjects received the instructions for the game before the rTMS train, it was possible to begin with the task about 30 s after completion of the stimulation train, thus under the influence of the rTMS aftereffect.

A.2 Instructions for Participants

We administered four different instructions, tailored to the participants' roles (investor or trustee) and to the information condition (anonymous or reputation). Instructions were identical across stimulation conditions, i.e., a trustee in the anonymous information condition was administered the same instructions regardless whether he received rDLPFC, lDLPFC or sham stimulation. Note that these instructions are adapted versions from the instructions to experiment 2, which was conducted prior to experiment 1. In the following, all four instructions are shown in the original layout, translated into English. The original German versions can be obtained from the author upon request.

A.2.1 Investor instructions, anonymous condition

General instructions for participants

You are now participating in an economic experiment that is funded by a number of research foundations.

Please read these instructions carefully. They will explain everything to you that you need for participation in the experiment. If you have questions, please raise your hand. We will then answer any questions at your work place. In all other cases, communication between the participants is strictly prohibited during the entire experiment.

At the beginning of the experiments, all participants will receive an initial endowment of 10 Swiss francs[1]. Any points you earn during the experiment will be converted to francs at the end of the experiment. The following exchange rate applies:

1 point = 20 centimes

At the end of the experiment, you will receive the income you earned during the experiment plus the 10 Swiss francs initial endowment paid in cash.

The experiment

In this experiment, a participant A is always paired with a participant B. No participant will learn with whom he was in any group of two, i.e. all decisions will be made anonymously. Your role ("participant A" or "participant B") will be determined at the beginning of the experiment; you will retain this role for the duration of the entire experiment. You will not learn of the identity of the participants assigned to you, neither before nor after the experiment. In the same way, the other participants will learn nothing about your identity.

The experiment consists of 15 periods; a different participant will be assigned to you in each period. This means that you will only meet each participant once at most.
Each period consists of two steps: in the first step, participant A transfers an amount of points to B. Participant B can transfer points back to A in the second step.

You are a participant A.

At the beginning of each period, each participant – A and B – receives an endowment of 10 points.

Your decision

You can decide how many points you would like to transfer to participant B. The experimenter will quadruple each point you transfer to participant B. If, for example, you transfer 1 point to participant B, participant B will receive 4 points; if you transfer 7 points to participant B, then participant B will receive 7 x 4 = 28 points.

[1] Translator's note: CHF 10.00 corresponds to $ 9.43, situation September 1, 2009.

The following intermediate point totals will thus result from your transfer:

Your point total: 10 – transfer

Participant B's point total: 10 + 4 × transfer

You must decide which of four possible amounts you would like to transfer to participant B: 1 point, 4 points, 7 points, or 10 points.

The following table summarizes the situation **after** you have made your decision and **before** B makes his or her decision.

You transfer...	Your point total	Participant B's point total
1 point	9	14
4 points	6	26
7 points	3	38
10 points	0	50

Participant B's decision:

In the second stage, participant B will learn how much you have transferred to him. Participant B can now decide how much he would like to return to you. He can choose between three different possibilities:

- Transfer **nothing**. The point totals then remain unchanged.
- Transfer **one-quarter** of the received amount. As he received four times the amount of your transfer, this corresponds to the amount that you transferred. In this case, your point total is again ten points, as it was at the beginning of the period.
- Transfer the sum back that gives both participants the **same number of points** (corresponds to a transfer of 62.5% of the points received).

The following table lists the earnings in points at the end of the period, after B has made his decision:

	Participant B's transfer:					
	"transfer nothing"		"Transfer one-quarter"		"Same number of points"	
Your transfer:	**You:**	**B:**	**You:**	**B:**	**You:**	**B:**
1 point	9	14	10	13	11.5	11.5
4 points	6	26	10	22	16	16
7 points	3	38	10	31	20.5	20.5
10 points	0	50	10	40	25	25

The period terminates after participant B has decided how many points he will transfer to you. A screen showing income will then inform you about the decisions made and the resulting income for this period. A new period with a new participant will then begin.

At the end of the experiment, your point income will be converted to francs and paid out to you in cash.

Do you have any questions?

Test questions

Please solve the following test questions. Your answers will have no consequence on your earnings; the questions only serve to see if all participants in the experiment have understood the rules. Please include all the steps of your calculations; this will aid us in finding any mistakes.

Question 1: Assume participant A transfers B 7 points. B then transfers A one-quarter of the sum he received. What are the income totals for this period?

A's point total: _____

B's point total: _____

Question 2: Assume participant A transfers B 1 point. B then transfers A the amount giving each the same number of points. What are the income totals for this period?

A's point total: _____

B's point total: _____

Question 3: Assume participant A transfers B 10 points. B then transfers A zero points. What are the income totals for this period?

A's point total: _____

B's point total: _____

Question 4: Assume participant A transfers B 10 points. B then transfers A the amount giving each the same number of points. What are the income totals for this period?

A's point total: _____

B's point total: _____

Please raise your hand once you have solved the test questions.

A.2.2 Investor instructions, reputation condition

General instructions for participants

You are now participating in an economic experiment that is funded by a number of research foundations.

Please read these instructions carefully. They will explain everything to you that you need for participation in the experiment. If you have questions, please raise your hand. We will then answer any questions at your work place. In all other cases, communication between the participants is strictly prohibited during the entire experiment.

At the beginning of the experiments, all participants will receive an initial endowment of 10 Swiss francs[1]. Any points you earn during the experiment will be converted to francs at the end of the experiment. The following exchange rate applies:

1 point = 20 centimes

At the end of the experiment, you will receive the income you earned during the experiment plus the 10 Swiss francs initial endowment paid in cash.

The experiment

In this experiment, a participant A is always paired with a participant B. No participant will learn with whom he was in any group of two, i.e. all decisions will be made anonymously. Your role ("participant A" or "participant B") will be determined at the beginning of the experiment; you will retain this role for the duration of the entire experiment. You will not learn of the identity of the participants assigned to you, neither before nor after the experiment. In the same way, the other participants will learn nothing about your identity.

The experiment consists of 15 periods; a different participant will be assigned to you in each period. This means that you will only meet each participant once at most.
Each period consists of two steps: in the first step, participant A transfers an amount of points to B. Participant B can transfer points back to A in the second step.

You are a participant A.

At the beginning of each period, each participant – A and B – receives an endowment of 10 points.

Your decision

You can decide how many points you would like to transfer to participant B. The experimenter will quadruple each point you transfer to participant B. If, for example, you transfer 1 point to participant B, participant B will receive 4 points; if you transfer 7 points to participant B, then participant B will receive 7 x 4 = 28 points.

[1] Translator's note: CHF 10.00 corresponds to $ 9.43, situation September 1, 2009.

The following intermediate point totals will thus result from your transfer:

Your point total: 10 – transfer

Participant B's point total: 10 + 4 × transfer

You must decide which of four possible amounts you would like to transfer to participant B: 1 point, 4 points, 7 points, or 10 points. **Before** you make your decision, you will be informed about the decisions Participant B made in the **last three** periods. We will explain this process below, after you have learned of Participant's B decision possibilities.

The following table summarizes the situation **after** you have made your decision and **before** B makes his or her decision.

You transfer...	Your point total	Participant B's point total
1 point	9	14
4 points	6	26
7 points	3	38
10 points	0	50

Participant B's decision:

In the second stage, participant B will learn how much you have transferred to him. Participant B can now decide how much he would like to return to you. He can choose between three different possibilities:

- Transfer **nothing**. The point totals then remain unchanged.
- Transfer **one-quarter** of the received amount. As he received four times the amount of your transfer, this corresponds to the amount that you transferred. In this case, your point total is again ten points, as it was at the beginning of the period.
- Transfer the sum back that gives both participants the **same number of points** (corresponds to a transfer of 62.5% of the points received).

The following table lists the earnings in points at the end of the period, after B has made his decision:

	Participant B's transfer:					
	"transfer nothing"		"Transfer one-quarter"		"Same number of points"	
Your transfer:	**You:**	**B:**	**You:**	**B:**	**You:**	**B:**
1 point	9	14	10	13	11.5	11.5
4 points	6	26	10	22	16	16
7 points	3	38	10	31	20.5	20.5
10 points	0	50	10	40	25	25

The period terminates after participant B has decided how many points he will transfer to you. A screen showing income will then inform you about the decisions made and the resulting income for this period. A new period with a new participant will then begin.

Before You determine your transfer to Participant B, you will be informed about the decisions Participant B made in the **three** previous periods. In particular, a list appears showing how many times B selected the options "transfer nothing", "transfer one-quarter", and "same number of points". If, for example, B opted for "transfer nothing" once, never chose "transfer one-quarter", and selected "same number of points" twice in the last three periods, the information for you will appear as follows.

Participant B made the following decisions in the last three periods:	
"transfer nothing"	1
"transfer one-quarter"	0
"same number of points"	2

When you decide how much you want to transfer to Participant B, you also know what he transferred back in the previous periods.

At the end of the experiment, your point income will be converted to francs and paid out to you in cash.

Do you have any questions?

Test questions

Please solve the following test questions. Your answers will have no consequence on your earnings; the questions only serve to see if all participants in the experiment have understood the rules. Please include all the steps of your calculations; this will aid us in finding any mistakes.

Question 1: Assume participant A transfers B 7 points. B then transfers A one-quarter of the sum he received. What are the income totals for this period?

A's point total: _____

B's point total: _____

Question 2: Assume participant A transfers B 1 point. B then transfers A the amount giving each the same number of points. What are the income totals for this period?

A's point total: _____

B's point total: _____

Question 3: Assume participant A transfers B 10 points. B then transfers A zero points. What are the income totals for this period?

A's point total: _____

B's point total: _____

Question 4: Assume participant A transfers B 10 points. B then transfers A the amount giving each the same number of points. What are the income totals for this period?

A's point total: _____

B's point total: _____

Question 5: Assume a Participant B is in period 3 and must make a decision for this period. In the previous two periods, Participant B selected "same number of points". If Participant B opts for "transfer nothing" in period 3, which information will the **next** Participant A receive about Participant B in next period 4? Please answer this question by completing the empty fields in this table as Participant A will see them in the next period.

Participant B made the following decisions in the last three periods:	
"transfer nothing"	_____
"transfer one-quarter"	_____
"same number of points"	_____

Please raise your hand once you have solved the test questions.

A.2.3 Trustee instructions, anonymous condition

General instructions for participants

You are now participating in an economic experiment that is funded by a number of research foundations.

Please read these instructions carefully. They will explain everything to you that you need for participation in the experiment. If you have questions, please raise your hand. We will then answer any questions at your work place. In all other cases, communication between the participants is strictly prohibited during the entire experiment.

At the beginning of the experiments, all participants will receive an initial endowment of 60 Swiss francs[1]. Any points you earn during the experiment will be converted to francs at the end of the experiment. The following exchange rate applies:

1 point = 10 centimes

At the end of the experiment, you will receive the income you earned during the experiment plus the 60 Swiss francs initial endowment paid in cash.

The experiment

In this experiment, a participant A is always paired with a participant B. No participant will learn with whom he was in any group of two, i.e. all decisions will be made anonymously. Your role ("participant A" or "participant B") will be determined at the beginning of the experiment; you will retain this role for the duration of the entire experiment. You will not learn of the identity of the participants assigned to you, neither before nor after the experiment. In the same way, the other participants will learn nothing about your identity.

The experiment consists of 15 periods; a different participant will be assigned to you in each period. This means that you will only meet each participant once at most.
Each period consists of two steps: in the first step, participant A transfers an amount of points to B. Participant B can transfer points back to A in the second step.

You are a participant B.

At the beginning of each period, each participant – A and B – receives an endowment of 10 points.

Participant A's decision

Participant A can decide how many points he would like to transfer to you. The experimenter will quadruple each point A transfers to you. If, for example, A transfers 1 point to you, you will receive 4 points; if A transfers 7 points to you, then you will receive 7 x 4 = 28 points.

[1] Translator's note: CHF 60.00 corresponds to $ 56.58, situation September 1, 2009.

The following intermediate point totals will thus result from your transfer:

Participant A's point total: 10 – transfer

Your point total: 10 + 4 × transfer

Participant A must decide which of four possible amounts he would like to transfer to you: 1 point, 4 points, 7 points, or 10 points.

The following table summarizes the situation **after** A has made his decision and **before** you make your decision.

A transfers...	Participant A's point total	Your point total
1 point	9	14
4 points	6	26
7 points	3	38
10 points	0	50

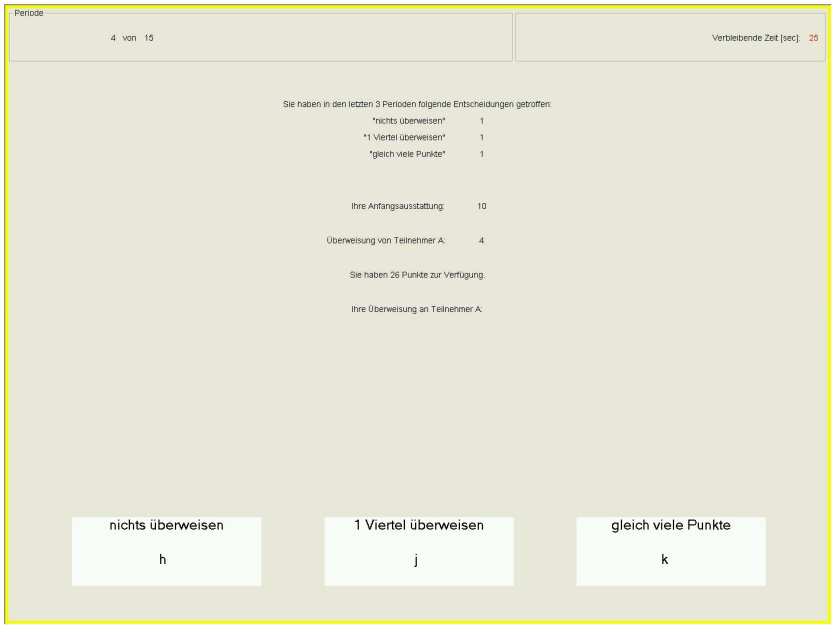
Your decision:

In the second stage, you will learn how much participant A has transferred to you. You can now decide how much you would like to return to participant A. You can choose between three different possibilities:

- Transfer **nothing**. The point totals then remain unchanged.
- Transfer **one-quarter** of the received amount. As you received four times the amount of A's transfer, this corresponds to the amount that A transferred. In this case, A's point total is again ten points, as it was at the beginning of the period.
- Transfer the sum back that gives both participants the **same number of points** (corresponds to a transfer of 62.5% of the points received).

The following table lists the earnings in points **at the end** of the period, after you have made your decision:

	Your transfer:					
	"transfer nothing"		**"Transfer one-quarter"**		**"Same number of points"**	
A's transfer:	**You:**	**B:**	**You:**	**B:**	**You:**	**B:**
1 point	9	14	10	13	11.5	11.5
4 points	6	26	10	22	16	16
7 points	3	38	10	31	20.5	20.5
10 points	0	50	10	40	25	25



Picture: Your decision screen

Use the laptop keys indicated above to enter your decision:

- h: "transfer nothing"
- j: "transfer one-quarter"
- k: "same number of points"

The period terminates after you have decided how many points you will transfer to Participant A. A screen showing income will then inform you about the decisions made and the resulting income for this period. A new period with a new participant will then begin.

You will also receive information on your decision screen showing which decisions you made in the last three periods. If, for example, you opted for "transfer nothing" once, never chose "transfer one-quarter", and selected "same number of points" twice, the information will appear as follows.

You made the following decisions in the last three periods:	
"transfer nothing"	1
"transfer one-quarter"	0
"same number of points"	2

At the end of the experiment, your point income will be converted to francs and paid out to you in cash.

Do you have any questions?

Test questions

Please solve the following test questions. Your answers will have no consequence on your earnings; the questions only serve to see if all participants in the experiment have understood the rules. Please include all the steps of your calculations; this will aid us in finding any mistakes.

Question 1: Assume participant A transfers B 7 points. B then transfers A one-quarter of the sum he received. What are the income totals for this period?
A's point total: _____
B's point total: _____

Question 2: Assume participant A transfers B 1 point. B then transfers A the amount giving each the same number of points. What are the income totals for this period?
A's point total: _____
B's point total: _____

Question 3: Assume participant A transfers B 10 points. B then transfers A zero points. What are the income totals for this period?
A's point total: _____
B's point total: _____

Question 4: Assume participant A transfers B 10 points. B then transfers A the amount giving each the same number of points. What are the income totals for this period?
A's point total: _____
B's point total: _____

Question 5: Assume you are in period 3 and must make a decision for this period. In the previous two periods, you selected "same number of points". If you opt for "transfer nothing" in period 3, which information will appear in period 4? Please answer this question by completing the empty fields in this table as you will see them in the next period.

You made the following decisions in the last three periods:	
"transfer nothing"	_____
"transfer one-quarter"	_____
"same number of points"	_____

Please raise your hand once you have solved the test questions.

A.2.4 Trustee instructions, reputation condition

General instructions for participants

You are now participating in an economic experiment that is funded by a number of research foundations.

Please read these instructions carefully. They will explain everything to you that you need for participation in the experiment. If you have questions, please raise your hand. We will then answer any questions at your work place. In all other cases, communication between the participants is strictly prohibited during the entire experiment.

At the beginning of the experiments, all participants will receive an initial endowment of 60 Swiss francs[1]. Any points you earn during the experiment will be converted to francs at the end of the experiment. The following exchange rate applies:

1 point = 10 centimes

At the end of the experiment, you will receive the income you earned during the experiment plus the 60 Swiss francs initial endowment paid in cash.

The experiment

In this experiment, a participant A is always paired with a participant B. No participant will learn with whom he was in any group of two, i.e. all decisions will be made anonymously. Your role ("participant A" or "participant B") will be determined at the beginning of the experiment; you will retain this role for the duration of the entire experiment. You will not learn of the identity of the participants assigned to you, neither before nor after the experiment. In the same way, the other participants will learn nothing about your identity.

The experiment consists of 15 periods; a different participant will be assigned to you in each period. This means that you will only meet each participant once at most.
Each period consists of two steps: in the first step, participant A transfers an amount of points to B. Participant B can transfer points back to A in the second step.

You are a participant B.

At the beginning of each period, each participant – A and B – receives an endowment of 10 points.

Participant A's decision

Participant A can decide how many points he would like to transfer to you. The experimenter will quadruple each point A transfers to you. If, for example, A transfers 1 point to you, you will receive 4 points; if A transfers 7 points to you, then you will receive 7 x 4 = 28 points.

[1] Translator's note: CHF 60.00 corresponds to $ 56.58, situation September 1, 2009.

The following intermediate point totals will thus result from your transfer:

Participant A's point total: 10 – transfer

Your point total: 10 + 4 × transfer

Participant A must decide which of four possible amounts he would like to transfer to you: 1 point, 4 points, 7 points, or 10 points. **Before** Participant A makes his decision, he will be informed about the decisions you made in the last **three** periods. We will explain this process below, after you have learned of your own decision possibilities.

The following table summarizes the situation **after** A has made his decision and **before** you make your decision.

A transfers...	Participant A's point total	Your point total
1 point	9	14
4 points	6	26
7 points	3	38
10 points	0	50

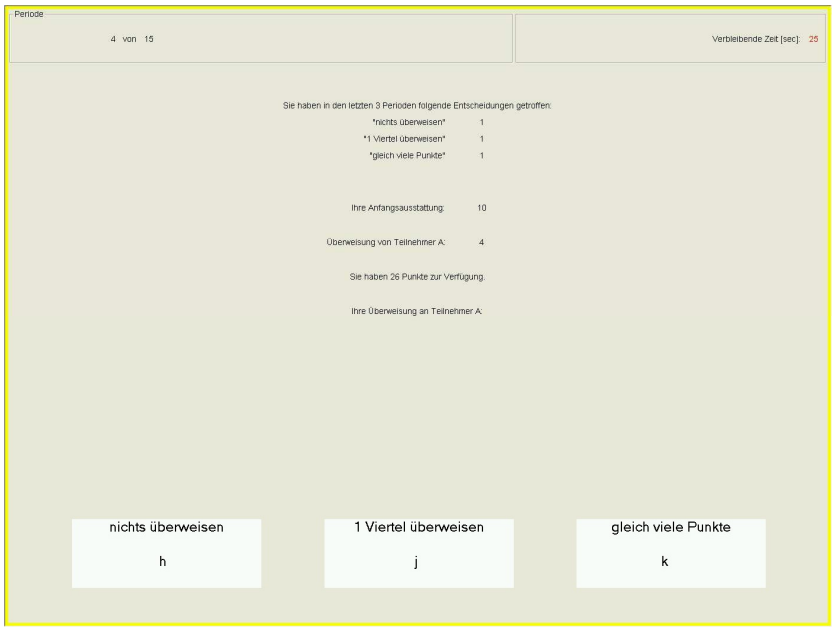
Your decision:

In the second stage, you will learn how much participant A has transferred to you. You can now decide how much you would like to return to participant A. You can choose between three different possibilities:

- Transfer **nothing**. The point totals then remain unchanged.
- Transfer **one-quarter** of the received amount. As you received four times the amount of A's transfer, this corresponds to the amount that A transferred. In this case, A's point total is again ten points, as it was at the beginning of the period.
- Transfer the sum back that gives both participants the **same number of points** (corresponds to a transfer of 62.5% of the points received).

The following table lists the earnings in points **at the end** of the period, after you have made your decision:

	Your transfer:					
	"transfer nothing"		**"Transfer one-quarter"**		**"Same number of points"**	
A's transfer:	**You:**	**B:**	**You:**	**B:**	**You:**	**B:**
1 point	9	14	10	13	11.5	11.5
4 points	6	26	10	22	16	16
7 points	3	38	10	31	20.5	20.5
10 points	0	50	10	40	25	25



Picture: Your decision screen

Use the laptop keys indicated above to enter your decision:

- h: "transfer nothing"
- j: "transfer one-quarter"
- k: "same number of points"

The period terminates after you have decided how many points you will transfer to Participant A. A screen showing income will then inform you about the decisions made and the resulting income for this period. A new period with a new participant will then begin.

Before Participant A determines his transfer to you, he will be informed about the decisions you as Participant B made in the **three** previous periods. In particular, a list appears showing how many times you selected the options "transfer nothing", "transfer one-quarter", and "same number of points" in the last three periods. If, for example, you opted for "transfer nothing" once, never chose "transfer one-quarter", and selected "same number of points" twice, the information for Participant A will appear as follows.

Participant B made the following decisions in the last three periods:	
"transfer nothing"	1
"transfer one-quarter"	0
"same number of points"	2

When Participant A decides how much he wants to transfer to you, he also knows what you transferred back in the previous periods.

You will also receive the information on the decisions you made in the last three periods.

At the end of the experiment, your point income will be converted to francs and paid out to you in cash.

Do you have any questions?

Test questions

Please solve the following test questions. Your answers will have no consequence on your earnings; the questions only serve to see if all participants in the experiment have understood the rules. Please include all the steps of your calculations; this will aid us in finding any mistakes.

Question 1: Assume participant A transfers B 7 points. B then transfers A one-quarter of the sum he received. What are the income totals for this period?
A's point total: _____
B's point total: _____

Question 2: Assume participant A transfers B 1 point. B then transfers A the amount giving each the same number of points. What are the income totals for this period?
A's point total: _____
B's point total: _____

Question 3: Assume participant A transfers B 10 points. B then transfers A zero points. What are the income totals for this period?
A's point total: _____
B's point total: _____

Question 4: Assume participant A transfers B 10 points. B then transfers A the amount giving each the same number of points. What are the income totals for this period?
A's point total: _____
B's point total: _____

Question 5: Assume you are in period 3 and must make a decision for this period. In the previous two periods, you selected "same number of points". If you opt for "transfer nothing" in period 3, which information will your **next** Participant A receive about you in period 4? Please answer this question by completing the empty fields in this table as Participant A will see them in the next period.

Participant B made the following decisions in the last three periods:	
"transfer nothing"	_____
"transfer one-quarter"	_____
"same number of points"	_____

Please raise your hand once you have solved the test questions.

A.3 Questionnaire

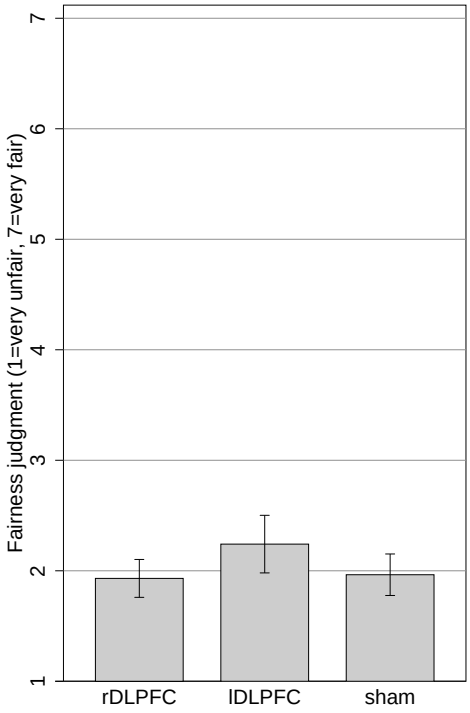
Measuring Subjects' Fairness Judgments. Because we hypothesized that rTMS to the right DLPFC would foster unfair behavior, we were interested in exploring whether rTMS also has a similar impact on the judgment of unfair behavior or whether the notion of fairness remains unchanged, thus creating a gap between judgment and choice. Directly after the completion of the experiment, subjects had to answer the following question: "Please indicate in the following how you evaluate the participant's behavior: Assume that a participant A has transferred 7 points to a participant B. Participant B then chose the option 'transfer nothing.' How do you evaluate the fairness of participant B's behavior?" Responses to this question were given on a 7-point Likert scale ranging from 1 ("very unfair") to 7 ("very fair").

Measuring Subjects' Ability to Assess the Future Consequences of Past Back Transfers. One alternative explanation for our findings could be that rTMS of the right DLPFC does not remove the ability to override immediate short-run benefits, but simply affects subjects' ability to assess the future consequences of previous back transfer behavior. To examine this possibility, we asked subjects how many points (1, 4, 7, or 10) they would expect an investor to transfer to a trustee who had opted twice for "equalize payoffs" and once for transferring back nothing in the previous 3 periods. The exact wording of the question was: "Assume that in the previous 3 periods, participant B has chosen once not to transfer anything, and has chosen twice to 'equalize payoffs.' How much do you expect the investor with whom B is matched next to transfer to B?" Subjects' response options were 1, 4, 7, or 10 points.

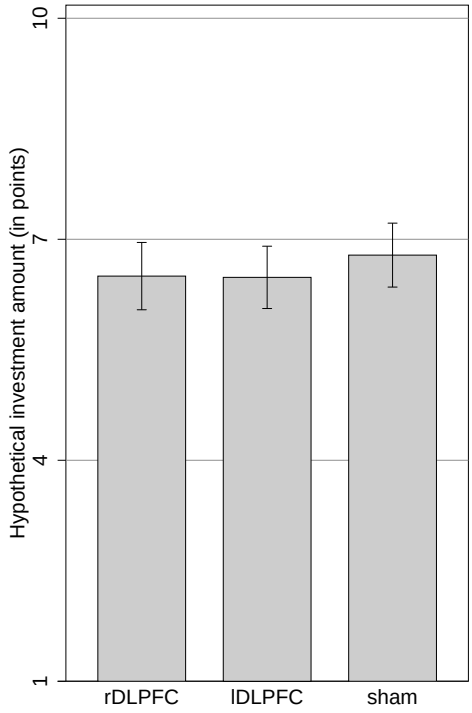
Fairness Judgment and Assessment of Future Consequences Across Treatments. As mentioned in chapter 1, we found no significant differences across stimulation conditions in the answers to the two questions posed earlier. The bar graphs in Figure A.1 illustrate the mean responses to the

Figure A.1: Trustees' responses to the two questions administered immediately after the trust game. Error bars show standard error of the mean.

(a) Trustees' fairness ratings.



(b) Trustees' assessment of the future consequences of a profile of back transfers.



fairness judgment question and the hypothetical investment question by stimulation condition. Figure A.1a shows how subjects assessed the fairness of a hypothetical trustee who returns nothing in response to an investor transfer of 7 points on a 7-point Likert scale. Means (all $p > 0.16$, Mann-Whitney U tests) and distributions ($p = 0.38$, Kruskal-Wallis test) do not differ significantly across the stimulation groups. Figure A.1b shows how subjects predicted the likely transfer of a hypothetical investor who observes that his current trustee opted twice for "equalize payoffs" and once for transferring back nothing in the previous 3 periods. Again, both means (all $p > 0.78$,

Mann-Whitney tests) and distributions ($p = 0.95$, Kruskal-Wallis test) do not differ significantly across the stimulation groups.

Measuring Subjects' Impulsivity and Personal Norm of Reciprocity. Approximately 10 days after the experiment, we sent the trustees a questionnaire that included the BIS and BAS scales developed by Carver and White (1994) (24 items) and translated into German by Strobel et al. (2001). This inventory investigates a subject's impulsive reaction to aversive stimuli (BIS) and rewarding stimuli (BAS). The questionnaire also included the 27-item Personal Norm of Reciprocity scale of Perugini et al. (2003). Items in this inventory explore a subject's tendency to reward another person's positive behavior (positive reciprocity) and to punish negative behavior (negative reciprocity). This questionnaire was translated at our institute and checked by back-translation.

A.4 Regression Analysis

To investigate whether the decision of how much to transfer back differed across stimulation conditions, we used GLS regression models. In the regression models, the dependent variable is the fraction of received points that the trustee transfers back. The dummy variables "right DLPFC" and "left DLPFC" are included to model the baseline effect of the 3 stimulation conditions (condition "sham" is the omitted category). The temptation to defect—and thus the self-control effort required for reputation formation—is largest if the investor transfers 10 points to the trustee. Thus, we hypothesize that the required recruitment of right DLPFC is highest in this case, implying that disruption of right DLPFC function is more likely to generate a behavioral effect. For this reason, our regressions also include a dummy variable for all those observations in which the investor actually transferred 10 points. The differences in the effect of stimulation conditions on the back transfer decision (for all situations in which the investor was sending 10 points) are

then estimated by including the corresponding interaction terms. We interact each stimulation condition (right DLPFC, left DLPFC, and sham) with a dummy variable indicating an investment of 10 points. For example, the interaction term between right DLPFC and an investment of 10 (denoted "right DLPFC $\times$ investment = 10") identifies the difference in the effect on back transfers between the right DLPFC condition and the sham stimulation condition (i.e., the omitted category) for the investment = 10 case. Note that the incentive for reputation formation disappears in the last period because the interaction between investors and trustees does not continue after period 15. Thus, we dropped the last observed period (period 15) from the data for all experimental conditions and for all observed experimental participants.

Tables A.1 and A.2 give the results of GLS estimation for the anonymous condition and the reputation condition, respectively. We present linear regression models in these tables, but the results obtained by the corresponding ordered discrete choice models for the 3 choice categories of the investor are identical with the results given here. We use robust standard errors adjusted for clustering on the subject level. We also include a random effect—assumed to be normally distributed—for each subject in our sample. We estimated various specifications of the regression model for each experimental condition (reputation and anonymous), and report 4 specifications here. Specification 1 tests whether the effect of stimulation of the right DLPFC and the left DLPFC differs from the effect of sham stimulation because the variable "sham DLPFC $\times$ investment = 10" is the omitted category. In specification 2, the variable "left DLPFC $\times$ investment = 10" is the omitted category; this specification tests whether the effect of stimulation of the right DLPFC and sham stimulation differs from the effect of stimulation of the left DLPFC. In specifications 3 and 4 we also control for individual subjects' fairness judgments, impulsivity, and reciprocity norm.

In Table A.1, we see that the interaction term "right DLPFC $\times$ investment = 10" is highly significant and negative in all 4 specifications. Specification 1 shows that, controlling for baseline differences captured by the variables right

DLPFC and left DLPFC, in the investment = 10 case, subjects whose right DLPFC was stimulated transferred back 11.4 percentage points less than subjects in the sham stimulation condition. Similarly, specification 2 shows that subjects in the right DLPFC condition transferred back 19.8 percentage points less than subjects in the left DLPFC stimulation condition. This indicates that, regardless of whether we compare the effect of stimulation of the right DLPFC with either the sham stimulation or stimulation of the left DLPFC, stimulation of the right DLPFC had a significantly negative effect on the trustees' back transfers. However, the effects of left DLPFC and sham stimulation on back transfers did not differ significantly from each other, as demonstrated by the insignificant coefficient estimates of the dummy variables "left DLPFC × investment = 10" and "sham × investment = 10."

As noted above, specifications 3 and 4 include further control variables and reveal that those who judged the scenario described earlier fairer than others tended to transfer back less. Including the control variables does not mitigate the strength of the effect of right DLPFC stimulation on back transfers.

Table A.2 shows the same 4 specifications as in Table A.1 for data from the anonymous condition. We see no significant differences in the effects of stimulation of the right DLPFC, stimulation of the left DLPFC, and sham stimulation on back transfers, because the estimated coefficients of the variables "right DLPFC × investment = 10," "left DLPFC × investment = 10," and "sham DLPFC × investment = 10" are always insignificant.

The estimations reported in Table A.1 and A.2 indicate a significant differential effect of rTMS across stimulations (right DLPFC, left DLPFC, sham) in the reputation condition, but not in the anonymous condition. Does the strength of this differential effect across stimulations also differ significantly between the two experimental conditions, reputation and anonymous? To investigate this question, we pooled all data from both experimental conditions (reputation and anonymous). We used the same four regression specifications as in Table A.1 and A.2, and also interacted all variables with a dummy vari-

able, "reputation," that indicates the experimental condition (i.e., whether or not an observation stems from the reputation condition). We report the corresponding estimation results in Table A.3. We see that the interaction of the factors "reputation × right DLPFC × investment = 10" is significant and negative in all specifications, regardless of which control variables we include ($p = 0.01$ for differences in right DLPFC vs. sham stimulation effects between experimental conditions in both specifications 1 and 3; $p < 0.01$ for differences in right DLPFC vs. left DLPFC stimulation effects between experimental conditions in both specifications 2 and 4). In contrast, the interactions for the other stimulations "reputation × left DLPFC × investment = 10" and "reputation × sham × investment = 10" are insignificant in all specifications. Finally, note that the coefficient of the dummy variable "reputation" is always positive (0.25 in both specifications 1 and 2, and 0.22 in both specifications 3 and 4) and always significant. This indicates that average back transfers in the reputation condition were 22 to 25 percentage points higher in the reputation condition than in the anonymous condition; that is, trustees cared greatly about their reputation when reputation formation was possible.

Table A.1: Reputation condition: GLS regression of the back transfer decisions on indicators of the experimental conditions and other controls

Dependent variable: relative back transfer	(1)	(2)	(3)	(4)
Right DLPFC	−0.00295	−0.00295	−0.00789	−0.00789
	(0.03479)	(0.03479)	(0.03565)	(0.03565)
Left DLPFC	−0.02785	−0.02785	0.00413	0.00413
	(0.03697)	(0.03697)	(0.03714)	(0.03714)
Investment=10	−0.10216**	−0.01789	−0.10732**	−0.03952
	(0.03321)	(0.03352)	(0.03334)	(0.03305)
rDLPFC × Inv.=10	−0.11380**	−0.19807**	−0.11095**	−0.17875**
	(0.04669)	(0.04692)	(0.04646)	(0.04644)
lDLPFC × Inv.=10	0.08427		0.06780	
	(0.04719)		(0.04674)	
sham × Inv.=10	0.08427	−0.08427	0.06780	−0.06780
	(0.04719)	(0.04719)	(0.04674)	(0.04674)
Fairness Judgment			−.06128**	−0.06128**
			(0.01637)	(0.01637)
Impulsivity (BIS)			−0.00864	−0.00864
			(0.04164)	(0.04164)
Impulsivity (BAS)			0.00629	0.00629
			(0.04383)	(0.04383)
Reciprocity (Positive)			0.04148	0.04148
			(0.02186)	(0.02186)
Reciprocity (Negative)			0.01750	0.01750
			(0.01846)	(0.01846)
Constant	0.49751**	0.49751**	0.35583*	0.35583*
	(0.02550)	(0.02550)	(0.16407)	(0.16407)
Observations	616/44	616/44	616/44	616/44
R-squared (within)	0.13	0.13	0.13	0.13

Standard errors, clustered on subject level, in parentheses.
* p<0.05, ** p<0.02

Table A.2: Anonymous condition: GLS regression of the back transfer decisions on indicators of the experimental conditions and other controls

Dependent variable: relative back transfer	(1)	(2)	(3)	(4)
Right DLPFC	−0.04169	−.04169	−0.08184	−0.08184
	(0.05937)	(0.05937)	(0.06275)	(0.06275)
Left DLPFC	−0.02250	−0.02250	−0.03246	−0.03246
	(0.06082)	(0.06082)	(0.06403)	(0.06403)
Investment=10	0.00446	0.05210	0.00564	0.054225
	(0.03857)	(0.03543)	(0.03879)	(0.04714)
rDLPFC × Inv.=10	0.05848	0.01085	0.06325	0.014666
	(0.04898)	(0.04655)	(0.04980)	(0.04714)
lDLPFC × Inv.=10	0.04764		0.04859	
	(0.05237)		(0.05237)	
sham × Inv.=10		−0.04764	−0.04859	−0.04859
		(0.05237)	(0.05237)	(0.05237)
Fairness Judgment			−0.03649**	−0.03649**
			(0.01884)	(0.01884)
Impulsivity (BIS)			0.02246	0.02246
			(0.07246)	(0.07246)
Impulsivity (BAS)			0.07697	-0.07697
			(0.06358)	(0.06358)
Reciprocity (Positive)			−0.00427	0.00427
			(0.04176)	(0.04176)
Reciprocity (Negative)			−0.00770	-0.00770
			(0.04579)	(0.04579)
Constant	0.25549**	0.25549**	0.50856	0.50856
	(0.04576)	(.04576)	(0.35623)	(0.35623)
Observations	602/43	602/43	588/42	588/42
R-squared (within)	0.01	0.01	0.01	0.01

Standard errors, clustered on subject level, in parentheses.
* p<0.05, ** p<0.02

Table A.3: Pooled GLS regression of the back transfer decisions on indicators of the experimental conditions and other controls

Dependent variable: relative back transfer	(1)	(2)	(3)	(4)
Right DLPFC	−0.04243	−0.04243	−0.07975	−0.07975
	(0.05028)	(0.05028)	(0.05026)	(0.05026)
Left DLPFC	−0.02350	−0.02350	−0.03276	−0.03276
	(0.05130)	(0.05130)	(0.05126)	(0.05126)
Investment=10	0.00156	0.00156	0.00259	0.00259
	(0.03884)	(0.03884)	(0.03893)	(0.03893)
rDLPFC×Investment=10	0.06078	0.06078	0.06490	0.06490
	(0.04971)	(0.04971)	(0.05051)	(0.05051)
lDLPFC×Investment=10	0.05055	0.05055	0.05377	0.05377
	(0.05291)	(0.05291)	(0.05269)	(0.05269)
Reputation	0.25154**	0.25154**	0.22103**	0.22103**
	(0.05115)	(0.05115)	(0.05028)	(0.05028)
Reput.×rDLPFC	0.03685	0.03685	0.07298	0.07298
	(0.06810)	(0.06810)	(0.06753)	(0.06753)
Reput.×lDLPFC	0.00145	0.00145	0.03424	0.03424
	(0.07000)	(0.07000)	(0.06919)	(0.06919)
Reput.×Inv.=10	−0.12475**	−0.10497*	−0.12663**	−0.11446**
	(0.05101)	(0.04877)	(0.05100)	(0.04848)
Reput.×rDLPFC×Inv.=10	−0.17306**	−0.19284**	−0.17430**	−0.18647**
	(0.06753)	(0.06585)	(0.06804)	(0.06617)
Reput.×lDLPFC×Inv.=10	0.01978		0.01217	
	(0.07057)		(0.07029)	
Reput.×sham×Inv.=10		−0.01978		−0.01217
		(0.07057)		(0.07029)
Fairness Judgment			−0.04238**	−0.04238**
			(0.01227)	(0.01227)
Impulsivity (BIS)			−0.00041	−0.00041
			(0.04031)	(0.04031)
Impulsivity (BAS)			−0.04067	−0.04067
			(0.03903)	(0.03903)
Reciprocity (Positive)			0.02919	0.02919
			(0.02196)	(0.02196)
Reciprocity (Negative)			0.01408	0.01408
			(0.01993)	(0.01993)
Constant	0.25648	0.25648	0.27945	0.27945
	(0.03852)	(0.03852)	(0.17066)	(0.17066)
Observations	1,218/87	1,218/87	1,204/86	1,204/86
R-squared (within)	0.07	0.07	0.08	0.08

Standard errors, clustered on subject level, in parentheses.
* p<0.05, ** p<0.02

Appendix B

Appendix to Experiment 2

B.1 Instructions for Participants

Instructions for the "anonymous" and the "implicit reputation" condition were identical. The only difference between the two conditions was the presence of implicit cues on the trustee screen. Thus, there are four kinds of instructions: investor instructions in the "anonymous"/"implicit reputation" condition; investor instructions in the "explicit reputation" condition; trustee instructions in the the "anonymous"/"implicit reputation" condition; trustee instructions for the "explicit reputation" condition.[1]

[1]Note that the instructions for experiment 1 are based on the instructions presented here.

B.1.1 Investor instructions, anonymous/implicit reputation condition

General instructions for participants

You are now participating in an economic experiment that is funded by a number of research foundations.

Please read these instructions carefully. They will explain everything to you that you need for participation in the experiment. If you have questions, please raise your hand. We will then answer any questions at your work place. In all other cases, communication between the participants is strictly prohibited during the entire experiment.

At the beginning of the experiments, all participants will receive an initial endowment of 10 Swiss francs[1]. Any points you earn during the experiment will be converted to francs at the end of the experiment. The following exchange rate applies:

1 point = 20 centimes

At the end of the experiment, you will receive the income you earned during the experiment plus the 10 Swiss francs initial endowment paid in cash.

The experiment

In this experiment, a participant A is always paired with a participant B. No participant will learn with whom he was in any group of two, i.e. all decisions will be made anonymously. Your role ("participant A" or "participant B") will be determined at the beginning of the experiment; you will retain this role for the duration of the entire experiment.
The experiment consists of 10 periods; a different participant will be assigned to you in each period. This means that you will only meet each participant once at most.
Each period consists of two steps: in the first step, participant A transfers an amount of points to B. Participant B can transfer points back to A in the second step.

You are a participant A.

At the beginning of each period, each participant – A and B – receives an endowment of 10 points.

Your decision

You can decide how many points you would like to transfer to participant B. The experimenter will quadruple each point you transfer to participant B. If, for example, you transfer 1 point to participant B, participant B will receive 4 points; if you transfer 7 points to participant B, then participant B will receive 7 x 4 = 28 points.

[1] Translator's note: CHF 10.00 corresponds to $ 9.43, situation September 1, 2009.

The following intermediate point totals will thus result from your transfer:

Your point total: 10 – transfer

Participant B's point total: 10 + 4 × transfer

You must decide which of four possible amounts you would like to transfer to participant B: 1 point, 4 points, 7 points, or 10 points.

The following table summarizes the situation **after** you have made your decision and **before** B makes his or her decision.

You transfer...	**Your point total**	**Participant B's point total**
1 point	9	14
4 points	6	26
7 points	3	38
10 points	0	50

Participant B's decision:

In the second stage, participant B will learn how much you have transferred to him. Participant B can now decide how much he would like to return to you. He can choose between three different possibilities:

- Transfer **nothing**. The point totals then remain unchanged.
- Transfer **one-quarter** of the received amount. As he received four times the amount of your transfer, this corresponds to the amount that you transferred. In this case, your point total is again ten points, as it was at the beginning of the period.
- Transfer the sum back that gives both participants the **same number of points** (corresponds to a transfer of 62.5% of the points received).

The following table lists the earnings in points **at the end** of the period, after B has made his decision:

	Participant B's transfer:					
	"transfer nothing"		**"Transfer one-quarter"**		**"Same number of points"**	
Your transfer:	**You:**	**B:**	**You:**	**B:**	**You:**	**B:**
1 point	9	14	10	13	11.5	11.5
4 points	6	26	10	22	16	16
7 points	3	38	10	31	20.5	20.5
10 points	0	50	10	40	25	25

The period terminates after participant B has decided how many points he will transfer to you. A screen showing income will then inform you about the decisions made and the resulting income for this period. A new period with a new participant will then begin.

At the end of the experiment, your point income will be converted to francs and paid out to you in cash.

Do you have any questions?

Test questions

Please solve the following test questions. Your answers will have no consequence on your earnings; the questions only serve to see if all participants in the experiment have understood the rules. Please include all the steps of your calculations; this will aid us in finding any mistakes.

Question 1: Assume participant A transfers B 7 points. B then transfers A one-quarter of the sum he received. What are the income totals for this period?

A's point total: _____

B's point total: _____

Question 2: Assume participant A transfers B 1 point. B then transfers A the amount giving each the same number of points. What are the income totals for this period?

A's point total: _____

B's point total: _____

Question 3: Assume participant A transfers B 10 points. B then transfers A zero points. What are the income totals for this period?

A's point total: _____

B's point total: _____

Question 4: Assume participant A transfers B 10 points. B then transfers A the amount giving each the same number of points. What are the income totals for this period?

A's point total: _____

B's point total: _____

Please raise your hand once you have solved the test questions.

B.1.2 Investor instructions, explicit reputation condition

General instructions for participants

You are now participating in an economic experiment that is funded by a number of research foundations.

Please read these instructions carefully. They will explain everything to you that you need for participation in the experiment. If you have questions, please raise your hand. We will then answer any questions at your work place. In all other cases, communication between the participants is strictly prohibited during the entire experiment.

At the beginning of the experiments, all participants will receive an initial endowment of 10 Swiss francs[1]. Any points you earn during the experiment will be converted to francs at the end of the experiment. The following exchange rate applies:

1 point = 20 centimes

At the end of the experiment, you will receive the income you earned during the experiment plus the 10 Swiss francs initial endowment paid in cash.

The experiment

In this experiment, a participant A is always paired with a participant B. No participant will learn with whom he was in any group of two, i.e. all decisions will be made anonymously. Your role ("participant A" or "participant B") will be determined at the beginning of the experiment; you will retain this role for the duration of the entire experiment.
The experiment consists of 10 periods; a different participant will be assigned to you in each period. This means that you will only meet each participant once at most.
Each period consists of two steps: in the first step, participant A transfers an amount of points to B. Participant B can transfer points back to A in the second step.

You are a participant A.

At the beginning of each period, each participant – A and B – receives an endowment of 10 points.

Your decision

You can decide how many points you would like to transfer to participant B. The experimenter will quadruple each point you transfer to participant B. If, for example, you transfer 1 point to participant B, participant B will receive 4 points; if you transfer 7 points to participant B, then participant B will receive 7 x 4 = 28 points.

[1] Translator's note: CHF 10.00 corresponds to $ 9.43, situation September 1, 2009.

The following intermediate point totals will thus result from your transfer:

Your point total: 10 – transfer

Participant B's point total: 10 + 4 × transfer

You must decide which of four possible amounts you would like to transfer to participant B: 1 point, 4 points, 7 points, or 10 points. **Before** you make your decision, you will be informed about the decisions Participant B made in the previous periods. We will explain this process below, after you have learned of Participant's B decision possibilities.

The following table summarizes the situation **after** you have made your decision and **before** B makes his or her decision.

You transfer...	Your point total	Participant B's point total
1 point	9	14
4 points	6	26
7 points	3	38
10 points	0	50

Participant B's decision:

In the second stage, participant B will learn how much you have transferred to him. Participant B can now decide how much he would like to return to you. He can choose between three different possibilities:

- Transfer **nothing**. The point totals then remain unchanged.
- Transfer **one-quarter** of the received amount. As he received four times the amount of your transfer, this corresponds to the amount that you transferred. In this case, your point total is again ten points, as it was at the beginning of the period.
- Transfer the sum back that gives both participants the **same number of points** (corresponds to a transfer of 62.5% of the points received).

The following table lists the earnings in points **at the end** of the period, after B has made his decision:

	Participant B's transfer:					
	"transfer nothing"		"Transfer one-quarter"		"Same number of points"	
Your transfer:	**You:**	**B:**	**You:**	**B:**	**You:**	**B:**
1 point	9	14	10	13	11.5	11.5
4 points	6	26	10	22	16	16
7 points	3	38	10	31	20.5	20.5
10 points	0	50	10	40	25	25

A screen showing income will then inform you about the decisions made and the resulting income for this period. A new period with a new participant will then begin.

Before you determine your transfer to Participant B, you will be informed about the decisions Participant B made in the previous periods. In particular, a list appears showing how many times B selected the options "transfer nothing", "transfer one-quarter", and "same number of points". If, for example, B opted for "transfer nothing" once, twice "transfer one-quarter", and selected "same number of points" three times, the information for you will appear as follows.

Participant B has made 6 decisions so far:	
"transfer nothing"	1
"transfer one-quarter"	2
"same number of points"	3

When you decide how much you want to transfer to Participant B, you also know what he transferred back in the previous periods.

At the end of the experiment, your point income will be converted to francs and paid out to you in cash.

Do you have any questions?

Test questions

Please solve the following test questions. Your answers will have no consequence on your earnings; the questions only serve to see if all participants in the experiment have understood the rules. Please include all the steps of your calculations; this will aid us in finding any mistakes.

Question 1: Assume participant A transfers B 7 points. B then transfers A one-quarter of the sum he received. What are the income totals for this period?

A's point total: _____

B's point total: _____

Question 2: Assume participant A transfers B 1 point. B then transfers A the amount giving each the same number of points. What are the income totals for this period?

A's point total: _____

B's point total: _____

Question 3: Assume participant A transfers B 10 points. B then transfers A zero points. What are the income totals for this period?

A's point total: _____

B's point total: _____

Question 4: Assume participant A transfers B 10 points. B then transfers A the amount giving each the same number of points. What are the income totals for this period?

A's point total: _____

B's point total: _____

Question 5: Assume a Participant B is in period 3 and must make a decision for this period. In the previous two periods, Participant B selected "same number of points". If Participant B opts for "transfer nothing" in period 3, which information will the **next** Participant A receive about Participant B in next period 4? Please answer this question by completing the empty fields in this table as Participant A will see them in the next period.

Participant B has made __ decisions so far:	
"transfer nothing"	___
"transfer one-quarter"	___
"same number of points"	___

Please raise your hand once you have solved the test questions.

B.1.3 Trustee instructions, anonymous/implicit reputation condition

General instructions for participants

You are now participating in an economic experiment that is funded by a number of research foundations.

Please read these instructions carefully. They will explain everything to you that you need for participation in the experiment. If you have questions, please raise your hand. We will then answer any questions at your work place. In all other cases, communication between the participants is strictly prohibited during the entire experiment.

At the beginning of the experiments, all participants will receive an initial endowment of 10 Swiss francs[1]. Any points you earn during the experiment will be converted to francs at the end of the experiment. The following exchange rate applies:

1 point = 10 centimes

At the end of the experiment, you will receive the income you earned during the experiment plus the 10 Swiss francs initial endowment paid in cash.

The experiment

In this experiment, a participant A is always paired with a participant B. No participant will learn with whom he was in any group of two, i.e. all decisions will be made anonymously. Your role ("participant A" or "participant B") will be determined at the beginning of the experiment; you will retain this role for the duration of the entire experiment.

The experiment consists of 10 periods; a different participant will be assigned to you in each period. This means that you will only meet each participant once at most.
Each period consists of two steps: in the first step, participant A transfers an amount of points to B. Participant B can transfer points back to A in the second step.

You are a participant B.

At the beginning of each period, each participant – A and B – receives an endowment of 10 points.

Participant A's decision

Participant A can decide how many points he would like to transfer to you. The experimenter will quadruple each point A transfers to you. If, for example, A transfers 1 point to you, you will receive 4 points; if A transfers 7 points to you, then you will receive 7 x 4 = 28 points.

[1] Translator's note: CHF 10.00 corresponds to $ 9.43, situation September 1, 2009.

The following intermediate point totals will thus result from your transfer:

Participant A's point total: 10 – transfer

Your point total: 10 + 4 × transfer

Participant A must decide which of four possible amounts he would like to transfer to you: 1 point, 4 points, 7 points, or 10 points.

The following table summarizes the situation **after** A has made his decision and **before** you make your decision.

A transfers...	Participant A's point total	Your point total
1 point	9	14
4 points	6	26
7 points	3	38
10 points	0	50

Your decision:

In the second stage, you will learn how much participant A has transferred to you. You can now decide how much you would like to return to participant A. You can choose between three different possibilities:

- Transfer **nothing**. The point totals then remain unchanged.
- Transfer **one-quarter** of the received amount. As you received four times the amount of A's transfer, this corresponds to the amount that A transferred. In this case, A's point total is again ten points, as it was at the beginning of the period.
- Transfer the sum back that gives both participants the **same number of points** (corresponds to a transfer of 62.5% of the points received).

The following table lists the earnings in points **at the end** of the period, after you have made your decision:

	Your transfer:					
	"transfer nothing"		"transfer nothing"		"transfer nothing"	
A's transfer:	**A:**	**You:**	**A:**	**You:**	**A:**	**You:**
1 point	9	14	10	13	11.5	11.5
4 points	6	26	10	22	16	16
7 points	3	38	10	31	20.5	20.5
10 points	0	50	10	40	25	25

The period terminates after you have decided how many points you will transfer to Participant A. A screen showing income will then inform you about the decisions made and the resulting income for this period. A new period with a new participant will then begin.

You will also receive information on your decision screen showing which decisions you made in the previous periods. If, for example, you opted for "transfer nothing" once, chose "transfer one-quarter" twice, and selected "same number of points" three times, the information will appear as follows.

Participant B has made 6 decisions so far:	
"transfer nothing"	1
"transfer one-quarter"	2
"same number of points"	3

At the end of the experiment, your point income will be converted to francs and paid out to you in cash.

Do you have any questions?

Test questions

Please solve the following test questions. Your answers will have no consequence on your earnings; the questions only serve to see if all participants in the experiment have understood the rules. Please include all the steps of your calculations; this will aid us in finding any mistakes.

Question 1: Assume participant A transfers B 7 points. B then transfers A one-quarter of the sum he received. What are the income totals for this period?

A's point total: _____

B's point total: _____

Question 2: Assume participant A transfers B 1 point. B then transfers A the amount giving each the same number of points. What are the income totals for this period?

A's point total: _____

B's point total: _____

Question 3: Assume participant A transfers B 10 points. B then transfers A zero points. What are the income totals for this period?

A's point total: _____

B's point total: _____

Question 4: Assume participant A transfers B 10 points. B then transfers A the amount giving each the same number of points. What are the income totals for this period?

A's point total: _____

B's point total: _____

Question 5: Assume you are in period 3 and must make a decision for this period. In the previous two periods, you selected "same number of points". If you opt for "transfer nothing" in period 3, which information will appear in period 4? Please answer this question by completing the empty fields in this table as you will see them in the next period.

Participant B has made ___ decisions so far:	
"transfer nothing"	___
"transfer one-quarter"	___
"same number of points"	___

Please raise your hand once you have solved the test questions.

B.1.4 Trustee instructions, explicit reputation condition

General instructions for participants

You are now participating in an economic experiment that is funded by a number of research foundations.

Please read these instructions carefully. They will explain everything to you that you need for participation in the experiment. If you have questions, please raise your hand. We will then answer any questions at your work place. In all other cases, communication between the participants is strictly prohibited during the entire experiment.

At the beginning of the experiments, all participants will receive an initial endowment of 10 Swiss francs[1]. Any points you earn during the experiment will be converted to francs at the end of the experiment. The following exchange rate applies:

1 point = 10 centimes

At the end of the experiment, you will receive the income you earned during the experiment plus the 10 Swiss francs initial endowment paid in cash.

The experiment

In this experiment, a participant A is always paired with a participant B. No participant will learn with whom he was in any group of two, i.e. all decisions will be made anonymously. Your role ("participant A" or "participant B") will be determined at the beginning of the experiment; you will retain this role for the duration of the entire experiment.

The experiment consists of 10 periods; a different participant will be assigned to you in each period. This means that you will only meet each participant once at most.
Each period consists of two steps: in the first step, participant A transfers an amount of points to B. Participant B can transfer points back to A in the second step.

You are a participant B.

At the beginning of each period, each participant – A and B – receives an endowment of 10 points.

Participant A's decision

Participant A can decide how many points he would like to transfer to you. The experimenter will quadruple each point A transfers to you. If, for example, A transfers 1 point to you, you will receive 4 points; if A transfers 7 points to you, then you will receive 7 x 4 = 28 points.

[1] Translator's note: CHF 10.00 corresponds to $ 9.43, situation September 1, 2009.

The following intermediate point totals will thus result from your transfer:

Participant A's point total: 10 – transfer

Your point total: 10 + 4 × transfer

Participant A must decide which of four possible amounts he would like to transfer to you: 1 point, 4 points, 7 points, or 10 points. **Before** Participant A makes his decision, he will be informed about the decisions you made in the previous periods. We will explain this process below, after you have learned of your own decision possibilities.

The following table summarizes the situation **after** A has made his decision and **before** you make your decision.

A transfers...	Participant A's point total	Your point total
1 point	9	14
4 points	6	26
7 points	3	38
10 points	0	50

Your decision:

In the second stage, you will learn how much participant A has transferred to you. You can now decide how much you would like to return to participant A. You can choose between three different possibilities:

- Transfer **nothing**. The point totals then remain unchanged.
- Transfer **one-quarter** of the received amount. As you received four times the amount of A's transfer, this corresponds to the amount that A transferred. In this case, A's point total is again ten points, as it was at the beginning of the period.
- Transfer the sum back that gives both participants the **same number of points** (corresponds to a transfer of 62.5% of the points received).

The following table lists the earnings in points **at the end** of the period, after you have made your decision:

	Your transfer:					
	"transfer nothing"		"transfer nothing"		"transfer nothing"	
A's transfer:	**A:**	**You:**	**A:**	**You:**	**A:**	**You:**
1 point	9	14	10	13	11.5	11.5
4 points	6	26	10	22	16	16
7 points	3	38	10	31	20.5	20.5
10 points	0	50	10	40	25	25

The period terminates after you have decided how many points you will transfer to Participant A. A screen showing income will then inform you about the decisions made and the resulting income for this period. A new period with a new participant will then begin.

Before Participant A determines his transfer to you, he will be informed about the decisions you as Participant B made in the previous periods. In particular, a list appears showing how many times you selected the options "transfer nothing", "transfer one-quarter", and "same number of points". If, for example, you opted for "transfer nothing" once, chose "transfer one-quarter" twice, and selected "same number of points" three times, the information for Participant A will appear as follows.

Participant B has made 6 decisions so far:	
"transfer nothing"	1
"transfer one-quarter"	2
"same number of points"	3

When Participant A decides how much he wants to transfer to you, he also knows what you transferred back in the previous periods.

At the end of the experiment, your point income will be converted to francs and paid out to you in cash.

Do you have any questions?

Test questions

Please solve the following test questions. Your answers will have no consequence on your earnings; the questions only serve to see if all participants in the experiment have understood the rules. Please include all the steps of your calculations; this will aid us in finding any mistakes.

Question 1: Assume participant A transfers B 7 points. B then transfers A one-quarter of the sum he received. What are the income totals for this period?

A's point total: _____

B's point total: _____

Question 2: Assume participant A transfers B 1 point. B then transfers A the amount giving each the same number of points. What are the income totals for this period?

A's point total: _____

B's point total: _____

Question 3: Assume participant A transfers B 10 points. B then transfers A zero points. What are the income totals for this period?

A's point total: _____

B's point total: _____

Question 4: Assume participant A transfers B 10 points. B then transfers A the amount giving each the same number of points. What are the income totals for this period?

A's point total: _____

B's point total: _____

Question 5: Assume you are in period 3 and must make a decision for this period. In the previous two periods, you selected "same number of points". If you opt for "transfer nothing" in period 3, which information will your **next** Participant A receive about you in period 4? Please answer this question by completing the empty fields in this table as Participant A will see them in the next period.

Participant B has made ___ decisions so far:	
"transfer nothing"	___
"transfer one-quarter"	___
"same number of points"	___

Please raise your hand once you have solved the test questions.

Appendix C

Appendix to Experiment 3

C.1 Worker and Team Leader Instructions

Workers received instructions pertaining to their job. Team leaders received copies of the worker instructions. In addition, they received their own instructions with details about the experiment and rules for communicating with workers. In the following, team leader instructions are merged into worker instructions. Information accessible only to team leaders is marked by [TL]. Parts that are irrelevant for the experiment are omitted. A complete German version of the handbook can be obtained from the author upon request.

Front page [omitted]

1. The company [omitted]

2. The campaign [omitted]

3. Company products [omitted]

4. Promotion Procedure

Both hotspot and club shifts have a duration of 3 hours. There is no prolongation or abortion ahead of time. Hotspot shifts are from 5pm to 8pm and club shifts from 11pm to2am.

Each team leader (TL) takes care of two teams, each team consisting of two workers. The TL is the contact person for questions of any kind.

Preparation [TL]

A few days prior to the first shift, the TLs receive the schedules, material and clothing for all shifts.

List of material and clothing [omitted]:

Registration: TLs will enter the data collected during the promotion into our online database. For this purpose it is necessary that every TL registers on [web address omitted] and communicates his/her user name to our office for the TL account to be activated.

Division into wage groups [TL]

The University of Zürich is conducting a scientific study in the context of the promotion campaign.

Each team consists of a "worker 1" and a "worker 2". This classification is irrelevant for the function of the workers, it is identical for both! The classification has, however, an effect on the workers' wages.

During the first week (i.e., FRI and SAT in week 1) ALL workers (i.e., both "worker 1" and "worker 2") earn **12 €** instead of the declared 10 € per hour.

In the second week (i.e. FRI and SAT in week 2), teams are divided into **3** different groups.

Group HH: Both, "worker 1" and "worker 2" **continue** to earn **12 €** per hour (HH = "worker 1" High, "worker 2" High)

Group LL: Both, "worker 1" and "worker 2" NOW receive **9 €** per hour (LL = "worker 1" Low, "worker 2" Low)

Group HL: ONLY "worker 2" receives NOW **9 €** per hour, "worker 1" CONTINUES to receive **12 €** per hour (HL = "worker 1" High, "worker 2" Low)

HH	Week **1** (FRI / SAT)		Week **2** (FRI / SAT)
"worker 1":	**12** €/hour	→	**12** €/hour
"worker 2":	**12** €/hour	→	**12** €/hour

LL	Week **1** (FRI / SAT)		Week **2** (FRI / SAT)
"worker 1":	**12** €/hour	→	**9** €/hour
"worker 2":	**12** €/hour	→	**9** €/hour

H_T	Week **1** (FRI / SAT)		Week **2** (FRI / SAT)
"worker 1":	**12** €/hour	→	**12** €/hour
"worker 2":	**12** €/hour	→	**9** €/hour

In all TL documents, groups are marked with **HH**, **LL** and **HL**. Group classification of the teams is mentioned on the schedule.

Meeting at the beginning of the shift [TL]

The TL arranges the team's meeting point. The meetings with the different teams are spaced 15 minutes apart. (i.e., different teams do not meet!).

BEFORE EACH SHIFT: The TL prepares all material necessary for the shift. The TL distributes the material to the workers at the beginning of the shift and collects it again after the shift. The TL is responsible for the preparation of the preparation of the documents, in particular the worker protocol and the TL protocol (correct date, location, worker names, wage group). The TL takes this information from the schedule.

Club shifts: before the first shift, the TL distributes T-shirts which the workers keep until the end of the last shift.

To ensure that shifts start in time, the TL arranges meetings 20 to 30 minutes before the beginning of the shift. Potential specifics for the shift can be discussed during this meeting.

Announcement of changes in wages [TL]

Week **1**:

ALL groups, wage **increase** "All workers receive **12 €** instead of 10 € per hour. The manager has decided this."

Week **2**:

Group **HH**, **equal** wage "All workers continue to receive **12 €** per hour. The manager has decided this."

Group **LL**, wage **cut** for both "From now on, all workers receive **9 €** instead 12 € per hour. The manager has decided this."

Group **HL**, wage **cut** for "worker 2" "From now on, all "workers 2"receive **9 €** instead of 12 € per hour. The manager has decided this."

IMPORTANT: There are strict rules for the communication with the workers (including how to react to questions). See communication guideline below.

During the shift [TL]

The TL

1. is the first to be at the meeting point and prepares handover of material
2. defines the radius of action for each promoter
3. takes care that workers are spaced far apart
4. takes pictures of customers when possible
5. appears periodically at the points-of-sale
6. takes care of supplies for workers
7. takes care that points-of-sale are clean (no dumped flyers or tickets)
8. takes care of the time limit of the shift

End of the shift [TL]

The TL collects material, documents and cash from workers. The TL records the sold tickets. TL and workers fill in the worker protocol. Then, workers are allowed to leave.

AFTER THE WORKERS HAVE LEFT, the TL fills in the TL protocol.

Post-processing [TL]

After each shift, the TL enters collected data into the online database. TLs receive a briefing for this task. The worker protocols, TL protocols and lists are handed in at the company's central office after the last shift. Cash can be handed over earlier on receipt to the responsible managers. The same procedure is used in both the December and the January campaign.

5. Summary promotion task

With short and concise customer approaches, a sales talk is completed after 3 minutes, after completion of the customer information form. However, we calculate on an average of 4 to 5 minutes, resulting in 10 to 20 tickets per hour.

A shift lasts 3 hours. We target popular public places and clubs in order to find more potential customers than can be approached.

The task of each worker is to approach each person in the age range of 18 to 30 years and to inform him/her as concisely and targeted as possible about the promotional ticket, resulting in:

A set of two tickets

- sold for 5 € or
- registration via customer registration form

When the registration form has been filled in, it is collected by the worker, and the customer receives a flyer.
During the shift, the tickets are carefully stored and handed back to the TL after the shift. Then, the revenue is registered and given to the TL.

6. **Handling of clothing and material [omitted]**

7. DOs and DON'Ts [worker]

DOs

- ✓ Scheduled shifts are mandatory
- ✓ Appear on time at the meeting point
- ✓ In case of sudden indisposition, inform TL at least 6 hours before the shift
- ✓ Enthusiasm for the campaign and pleasure with the job
- ✓ Be cheerful and convincing
- ✓ Concise approaches
- ✓ Ensure readability and correctness of information. For this purpose, fill in the form for the customer and require an ID (tickes only for correct information, otherwise for 5 €)
- ✓ Careful handling of provided material
- ✓ Record worker initials on each customer form
- ✓ Arrange ride back independently. The TL is not responsible for that.

DON'Ts

- ✓ Cancel shifts on short notice
- ✓ Not being punctual
- ✓ Being in a bad mood and creating bad mood within the team
- ✓ Drinking alcohol during the shift
- ✓ Long cigarette breaks (5 minutes per hour max.)

Contact person

During the shift, your contact person is the TL responsible that for you. The TL also supplies you with all information pertaining to your job. For further inquiries, please contact the company's regional office at [address omitted].

8. Communication guideline [TL]

UNDER NO CIRCUMSTANCES mention to workers:

- ✓ Study of University of Zürich
- ✓ Existence of other wage groups
- ✓ To workers of the December campaign: Existence of the January campaign
- ✓ Subsequent verification of collected customer information

Emphasize in the presence of workers:

- ✓ Campaign is one-time, no further job opportunity
- ✓ Division in "worker 1" and "worker 2" was random and is irrelevant for the job
- ✓ The goal of 10-20 tickets per hour is non-binding and has no consequence for earnings

If workers ask:

- ✓ "Why are there a worker 1 and a worker 2?"
  Answer: *This was introduced for accounting purposes.*
- ✓ "Why does the wage change?"
  Answer: *I don't know, this is what the manager said.*
- ✓ "I have come to know that there are other wage groups."
  Answer: *I have no knowledge of that, I will enquire about this.*
- ✓ Other questions/remarks:
  Answer: *That's the way it is, it's what the manager said.*

9. Behavior guideline [TL]

- ✓ Treat all workers in the same way, no differential motivation or punishment.
- ✓ No active performance monitoring. The workers must not feel watched or controlled (e.g., when the TL comes to the point-of-sale and takes customer pictures).

10. Figures at a glance [omitted]

Please effectuate the instructions of this handbook, so that all goals can be attained and everything can be properly accounted for.

Thank you very much

C.2 A Model of Social Comparison

The standard economic model ignores fairness motives and social comparison because it assumes that individuals exclusively pursue their own material interest. Consequently, in the absence of reputation incentives, neither the level of the own flat wage nor the level of coworker wages has an impact on effort because in each case material work incentives are identical.

The model of inequity aversion developed by Fehr and Schmidt (1999), however, takes fairness and social comparison into account and provides a micro-foundation of the fair wage-effort hypothesis. In particular, the level of the flat wage determines firm and worker payoffs and thus influences inequity between the involved parties. An inequity averse worker therefore has an incentive to minimize this inequity by choosing a specific level of effort.

In our setup a firm employs two identical workers, worker 1 and worker 2. We analyze worker i's effort e_i in response to the wages w_i and w_j set by the firm, and to coworker effort e_j. Let the firm's payoff from worker i be the revenue generated by worker i minus wage cost:

$$\pi_i = ve_i - w_i, \quad i \in \{1,2\},\ e_i, w_i \geq 0.$$

Let worker i's payoff be her wage minus her effort cost:

$$x_i = w_i - ce_i, \quad 0 < c < v,$$

and let worker i's utility be the sum of her payoff x_i, her disutility from inequity with respect to the firm, and her disutility from inequity with respect to her coworker j:

$$\begin{aligned} U_i \;=\; & x_i - \frac{\alpha}{2}\max\{\pi_i - x_i, 0\} - \frac{\beta}{2}\max\{x_i - \pi_i, 0\} \\ & - \frac{\alpha}{2}\max\{x_j - x_i, 0\} - \frac{\beta}{2}\max\{x_i - x_j, 0\} \end{aligned}$$

with the assumption that $\beta \leq \alpha$ and $0 \leq \beta < 1$. The parameter α measures

how much the worker dislikes disadvantageous inequity and β measures how much she dislikes advantageous inequity.[1]

Inequity

Effort is not contractible, thus the only reason why a worker should exert effort is to reduce inequity. Consider first inequity between one worker and the firm:

$$|x_i - \pi_i| = |2w_i - (c+v)e_i|.$$

If the worker does not provide effort, i.e., $e_i = 0$, she creates inequity to her advantage of $x_i - \pi_i = 2w_i$. If she works, she will reduce her own payoff but at the same time increase the firm's payoff. One unit of effort decreases inequity with respect to the firm by $c + v$ units. For any wage w_i, the level of effort that equalizes worker and firm payoff is:

$$e_i^{x_i=\pi_i}(w_i) = \frac{2}{c+v} w_i \equiv e^F(w_i).$$

Consider now inequity between worker and coworker. Inequity in relation to the coworker depends on the wages and effort levels of both workers:

$$|x_i - x_j| = |w_i - w_j - c(e_i - e_j)|.$$

Hence, in contrast to worker-firm inequity, one unit of effort changes inequity in relation to the coworker by only c units because own effort does not affect coworker payoff. For wage levels w_i and w_j, and coworker effort e_j, the level of own effort that equalizes worker and coworker payoff is:

$$e_i^{x_i=x_j}(w_i, w_j, e_j) = e_j + \frac{w_i - w_j}{c} \equiv e^C.$$

In the special case of equal wages ($w_i = w_j$), worker payoffs are equalized if and only if they exert the same amount of effort.

[1] When $\alpha, \beta = 0$, the worker only cares about her own payoff, which corresponds to the standard economic model.

Effort Choice

We now analyze a worker's effort choice for different values of α and β. First, if a worker does not suffer enough from advantageous inequity (i.e., for β below a threshold $\underline{\beta}$), she will never exert effort. In other words, marginal utility from inequity reduction with respect to both firm and coworker is lower than marginal cost of effort:

$$\frac{\beta}{2}(c+v)+\frac{\beta}{2}<c, \quad x_i \geq \pi_i.$$

This is equivalent to a low value of β:

$$\beta < \frac{2c}{2c+v} \equiv \underline{\beta}.$$

Second, if a worker suffers much from advantageous inequity with respect to the firm, and not too much from disadvantageous inequity with respect to her coworker, she will always exert effort, no matter how much effort her coworker exerts:

$$\frac{\beta}{2}(c+v)-\frac{\alpha}{2}c>c, \quad x_i \geq \pi_i.$$

This corresponds to a low value of α and high value of β:

$$\alpha < \beta\frac{c+c}{v}-2 \equiv \bar{\alpha}$$

Hence, a worker with such inequity parameters will increase effort as long as her payoff is greater than firm payoff. However, no worker ever exerts more effort than the level that equalizes worker and firm payoff, i.e., $e_i = e^F$, even if thereby she could reduce advantageous inequity with respect to her coworker. This is because marginal cost from increased disadvantageous inequity with respect to the firm always outweighs marginal gain from inequity reduction with respect to the coworker:

$$-\frac{\alpha}{2}(c+v)+\frac{\beta}{2}c<0.$$

As a result, a worker with a low α and high β will always equalize payoffs with the firm by exerting e^F.

Third, if a worker suffers much from both advantageous inequity with respect to the firm and disadvantageous inequity with respect to her coworker, she would like to decrease inequity with respect to the firm because $\beta > \underline{\beta}$, but she is not willing to incur a lower payoff than her coworker because $\alpha > \bar{\alpha}$. Thus, if both α and β are high, she will always equalize payoffs with her coworker rather than with the firm by choosing effort level e^C.

Now we can characterize the three sets of Nash equilibrium strategies for worker 1 and worker 2 as a function of α and β:

$$\begin{array}{lll}
\text{(i)} & \beta < \underline{\beta}: & e_1^* = e_2^* = 0 \\
\text{(ii)} & \beta \geq \underline{\beta}, \alpha < \bar{\alpha}: & e_1^* = e^F(w_i), e_2^* = e^F(w_j) \\
\text{(iii)} & \beta \geq \underline{\beta}, \alpha \geq \bar{\alpha}: & e_1^* = e_1, e_2^* = e_1 + \dfrac{w_2 - w_1}{c}, \quad e_1 \in [0, e_1^F],\ e_2 \in [0, e_2^F]
\end{array}$$

Effort Choice Predictions across Treatments

Now suppose that the wage can take on two levels, H and L, with $H - L = \Delta > 0$. Consider the situation where both workers earn the high wage H. In case (i), workers provide zero effort in equilibrium. In case (ii), they provide positive effort $e^F(H) = \frac{2}{c+v}H$. In case (iii), any effort level between 0 and $e^F(H)$ that is chosen by both workers is a Nash equilibrium. However, if we assume that workers are able to coordinate on the coalition-proof Nash equilibrium, only two equilibria remain. These two equilibria again depend on the value of β. If β is below a threshold $\bar{\beta}$, workers care little about advantageous inequity with respect to the firm and coordinate on a effort level of 0. Conversely, if β is above $\bar{\beta}$, workers care much about inequity with respect to the firm and coordinate on $e^F(H)$:

$$\begin{array}{lll}
\text{(iii.a)} & \alpha \geq \bar{\alpha}, \underline{\beta} \leq \beta < \dfrac{2c}{c+v} \equiv \bar{\beta}: & e_1^* = e_2^* = 0 \\
\text{(iii.b)} & \alpha \geq \bar{\alpha}, \beta \geq \bar{\beta}: & e_1^* = e_2^* = e^F(H)
\end{array}$$

Figure C.1: Equilibrium effort levels in the four treatment groups

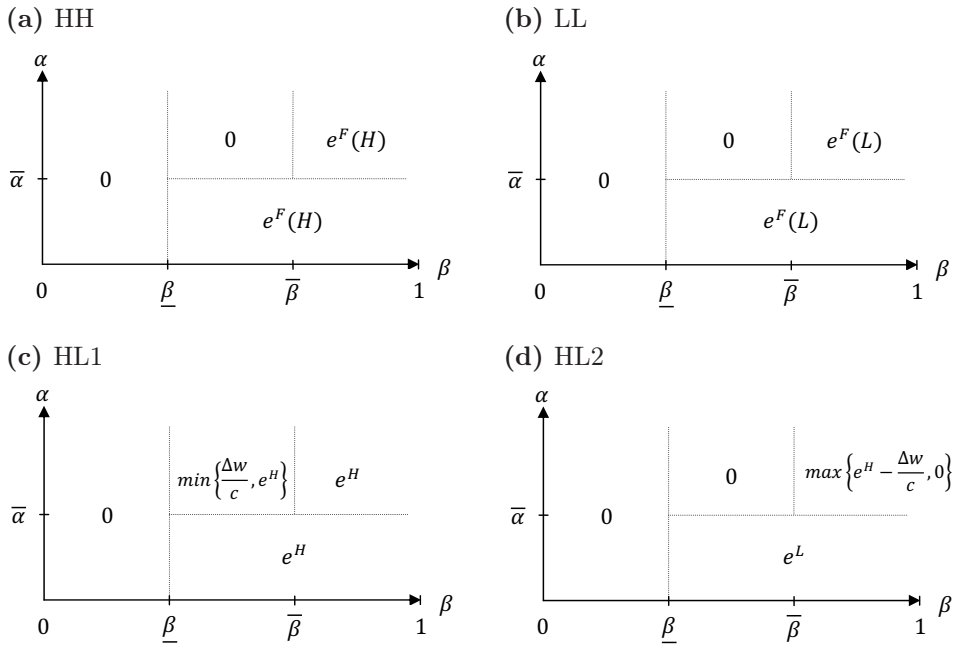
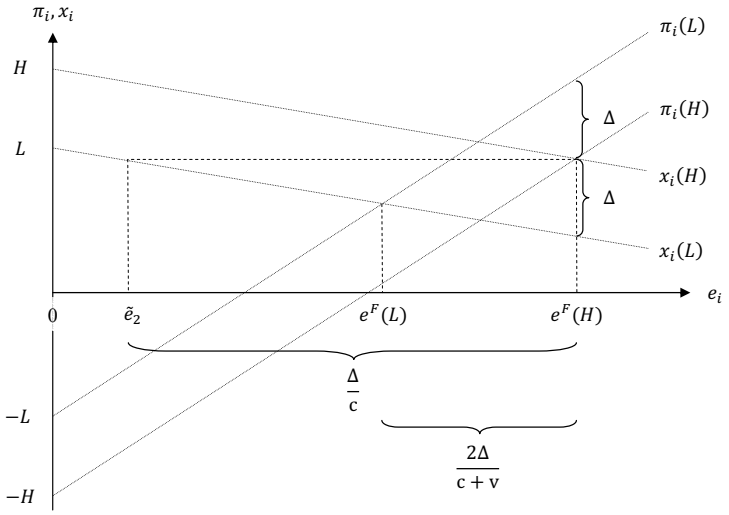


Figure C.1a shows the equilibrium effort levels if both workers earn the high wage.

The situation where both workers earn the low wage L is analogous to the situation above (see Figure C.1b). Thus, depending on the values of the inequity parameters, workers choose either $e_1^* = e_2^* = 0$ or $e_1^* = e_2^* = e^F(L)$. In the latter case, a lower wage implies a lower effort level, i.e., $e^F(L) = e^F(H) - \frac{2\Delta}{c+v} < e^F(H)$ because the firm-equalizing effort level $e^F = \frac{2}{c+v}w$ is proportional to the wage.

Now consider the situation where worker 1 earns the high wage H and worker 2 the low wage L (see Figures C.1c and C.1d). In case (i), where workers do not suffer much from inequity, equilibrium effort is 0 for both workers. In case (ii), where workers equalize their respective payoffs with the firm, equilibrium effort is $e^F(H)$ for worker 1, and $e^F(L)$ for worker 2. In case (iii), workers equalize payoffs with each other. In case (iii.a), worker 1 has

Figure C.2: Payoffs as functions of effort.



to choose a positive effort because worker 2 cannot provide negative effort. Worker 1 therefore chooses effort $\tilde{e}_1 = \frac{\Delta}{c}$ unless this value exceeds $e^F(H)$, in which case she chooses $e^F(H)$. In case (iii.b), worker 1 chooses $e^F(H)$. She is not willing to provide more effort than $e^F(H)$ because the utility gain from inequity reduction with respect to the coworker would always be lower than the associated effort cost. Thus, in order to equalize worker payoffs, worker 2 chooses $\tilde{e}_2 = e^F(H) - \frac{\Delta}{c}$ unless this value is negative, in which case she chooses 0.

Effort $\tilde{e}_2$ is lower than $e^F(L)$ because worker 2 can influence the payoff of the firm but not the payoff of her coworker. As can be seen in Figure C.2, to eliminate 2Δ units of inequity with respect to the firm, worker 2 has to reduce effort from $e^F(H)$ to $e^F(L)$. This reduction amounts to $\frac{2\Delta}{c+v}$ units because reducing effort by one unit not only increases own payoff by c units but also reduces firm payoff by v units. In contrast, to eliminate Δ units of inequity with respect to the coworker, worker 2 has to reduce effort by $\frac{\Delta}{c}$ units because reducing effort by one unit increases own payoff by c but leaves coworker payoff unaffected. Because marginal product of effort v is greater than marginal cost of effort c, $\tilde{e}_2$ is lower than $e^F(L)$.

Summary of Hypotheses

We now derive the hypotheses for the change in effort from pre- to post-intervention period across treatments:[2]

Hypothesis 1 (Treatment LL)

1a *If $\beta < \underline{\beta}$ or $\beta < \bar{\beta}, \alpha \geq \bar{\alpha}$: both workers' equilibrium effort levels remain the same.*

1b *Otherwise: both workers' equilibrium effort levels decrease.*

Hypothesis 2 (Treatment Group HL2)

2a *If $\beta < \underline{\beta}$ or $\beta < \bar{\beta}, \alpha \geq \bar{\alpha}$: worker 2's equilibrium effort level remains the same.*

2b *Otherwise: worker 2's equilibrium effort level decreases.*

2b' *If $\beta \geq \underline{\beta}, \alpha < \bar{\alpha}$: worker 2's equilibrium effort level is the same as in LL.*

2b" *If $\beta \geq \underline{\beta}, \alpha \geq \bar{\alpha}$: worker 2's equilibrium effort level is lower as in LL.*

Hypothesis 3 (Treatment Group HL1)

3a *If $\alpha < \bar{\alpha}$ or $\beta < \underline{\beta}$ or $\beta \geq \bar{\beta}, \alpha \geq \bar{\alpha}$: worker 1's equilibrium effort level remains the same.*

3b *Otherwise: worker 1's equilibrium effort level increases.*

[2]Treatment HH equilibrium effort levels are the same as pre-intervention period equilibrium effort levels because both workers continue to earn the high wage.

Bibliography

Abler, B., H. Walter, A. Wunderlich, J. Grothe, C. Schönfeldt-Lecuona, M. Spitzer, and U. Herwig (2005). Side effects of transcranial magnetic stimulation biased task performance in a cognitive neuroscience study. *Brain topography 17*(4), 193–196.

Adams, J. (1963). Toward an Understanding of Inequity. *Journal of Abnormal and Social Psychology 67*(5), 422–436.

Akerlof, G. and J. Yellen (1990). The fair wage-effort hypothesis and unemployment. *Quarterly Journal of Economics 105*(2), 255–283.

Allen, T. (2010, January 29). Euro area unemployment rate up to 10.0%. News Release 16/2010, Eurostat, Luxembourg.

Andreoni, J. and J. Miller (1993). Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence. *The Economic Journal 103*(418), 570–585.

Aron, A., T. Robbins, and R. Poldrack (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Sciences 8*(4), 170–177.

Bardsley, N. (2008). Dictator game giving: altruism or artefact? *Experimental Economics 11*(2), 122–133.

Basu, S., J. Dickhaut, G. Hecht, K. Towry, and G. Waymire (2009). Record-keeping alters economic history by promoting reciprocity. *Proceedings of the National Academy of Sciences 106*(4), 1009–1014.

Bateson, M., D. Nettle, and G. Roberts (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology Letters 12*, 412–414.

Berg, J., J. Dickhaut, and K. McCabe (1995). Trust, reciprocity, and social history. *Games and Economic Behavior 10*(1), 122–142.

Bernheim, B. and A. Rangel (2007). Behavioral public economics: Welfare and policy analysis with non-standard decision-makers. In P. Diamond and H. Vartiainen (Eds.), *Behavioral Economics and its Applications*, pp. 7–77. Princeton, NJ: Princeton University Press.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How Much Should We Trust Differences-in-Differences Estimates? *Quarterly Journal of Economics 119*(1), 249–275.

Bewley, T. (1999). *Why wages don't fall in a recession*. Cambridge, MA: Harvard University Press.

Bolton, G. and A. Ockenfels (2000). ERC: A theory of equity, reciprocity, and competition. *American Economic Review 90*(1), 166–193.

Bowles, S. and H. Gintis (2004). The evolution of strong reciprocity: cooperation in heterogeneous populations. *Theoretical Population Biology 65*(1), 17–28.

Boyd, R., H. Gintis, S. Bowles, and P. Richerson (2003). The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences 100*(6), 3531–3535.

Boyd, R. and P. Richerson (2005). *The origin and evolution of cultures*. New York City, NY: Oxford University Press, USA.

Brown, M., A. Falk, and E. Fehr (2004). Relational contracts and the nature of market interactions. *Econometrica 72*(3), 747–780.

Burnham, T. and B. Hare (2007). Engineering Human Cooperation. *Human Nature 18*(2), 88–108.

Camerer, C. (2003). *Behavioral game theory*. Princeton, NJ: Princeton University Press.

Camerer, C. and U. Malmendier (2005). Behavioral economics of organizations. In P. Diamond and H. Vartiainen (Eds.), *Behavioral Economics and its Applications*, pp. 235–281. Princeton, NJ: Princeton University Press.

Camerer, C. and K. Weigelt (1988). Experimental tests of a sequential equilibrium reputation model. *Econometrica 56*(1), 1–36.

Card, D. (1997). Deregulation and Labor Earnings in the Airline Industry. In J. Peoples (Ed.), *Regulatory reform and labor markets*, pp. 183–230. Norwell, MA: Kluwer Academic Publishers.

Carver, C. and T. White (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of personality and social psychology 67*, 319–333.

Charness, G. and P. Kuhn (2007). Does pay inequality affect worker effort? Experimental evidence. *Journal of Labor Economics 25*(4), 693–723.

Christie, R. and F. C. Geis (1970). *Studies in Machiavellianism*. New York: Academic Press.

Cochard, F., P. Nguyen Van, and M. Willinger (2004). Trusting behavior in a repeated investment game. *Journal of Economic Behavior and Organization 55*(1), 31–44.

Cohn, A., E. Fehr, and L. Goette (2009). Fair Wage and Effort - Evidence from a Field Experiment. Unpublished manuscript.

Cooper, D. and J. Kagel (2010). Other Regarding Preferences: A Selective Survey of Experimental Results. In J. H. Kagel and A. E. Roth (Eds.), *The Handbook of Experimental Economics*, Volume 2. Princeton, NJ: Princeton University Press.

Croson, R. (2007). Theories of commitment, altruism and reciprocity: Evidence from linear public goods games. *Economic Inquiry 45*(2), 199–216.

Davis, J. (1959). A formal interpretation of the theory of relative deprivation. *Sociometry 22*(4), 280–296.

Dawes, C., J. Fowler, T. Johnson, R. McElreath, and O. Smirnov (2007). Egalitarian motives in humans. *Nature 446*, 794–796.

Delgado, M., R. Frank, and E. Phelps (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nature Neuroscience 8*(11), 1611–1618.

Dufwenberg, M. and G. Kirchsteiger (2004). A theory of sequential reciprocity. *Games and Economic Behavior 47*(2), 268–298.

Eisenegger, C., V. Treyer, E. Fehr, and D. Knoch (2008). Time-course of "off-line" prefrontal rTMS effects—a PET study. *NeuroImage 42*(1), 379–384.

Engelmann, D. and U. Fischbacher (2009). Indirect reciprocity and strategic reputation building in an experimental helping game. *Games and Economic Behavior 67*(2), 399–407.

Falk, A. and U. Fischbacher (2006). A theory of reciprocity. *Games and Economic Behavior 54*(2), 293–315.

Falk, A., S. Gächter, and J. Kovács (1999). Intrinsic motivation and extrinsic incentives in a repeated game with incomplete contracts. *Journal of Economic Psychology 20*(3), 251–284.

Falk, A. and J. Heckman (2009). Lab Experiments Are a Major Source of Knowledge in the Social Sciences. *Science 326*(5952), 535–538.

Fehr, E. (2000). Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives 14*(3), 159–181.

Fehr, E., M. Brown, and C. Zehnder (2009). On Reputation: A Microfoundation of Contract Enforcement and Price Rigidity. *The Economic Journal 119*(536), 333–353.

Fehr, E. and A. Falk (1999). Wage rigidity in a competitive incomplete contract market. *Journal of Political Economy 107*(1), 106–134.

Fehr, E. and U. Fischbacher (2003). The nature of human altruism. *Nature 425*(6960), 785–791.

Fehr, E., U. Fischbacher, and S. Gächter (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature 13*(1), 1–25.

Fehr, E. and S. Gächter (2000). Cooperation and punishment in public goods experiments. *American Economic Review 90*(4), 980–994.

Fehr, E. and L. Götte (2007). Do Workers Work More if Wages Are High? Evidence from a randomized field experiment. *American Economic Review 97*(1), 298–317.

Fehr, E. and K. Schmidt (1999). A Theory of Fairness, Competition, and Cooperation. *Quarterly Journal of Economics 114*(3), 817–868.

Fehr, E. and F. Schneider (2009). Eyes are on us, but nobody cares: are eye cues relevant for strong reciprocity? *Proceedings of the Royal Society B: Biological Sciences*, –. published online ahead of print.

Festinger, L. (1954). A theory of social comparison processes. *Human relations 7*(2), 117–140.

Fischbacher, U. (2007). z-Tree: Zurich Toolbox for Ready-made Economic Experiments. *Experimental Economics 10*(2), 171–178.

Fischbacher, U., S. Gächter, and E. Fehr (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters 71*(3), 397–404.

Friedman, D. (1995). *The machinery of freedom: guide to a radical capitalism.* Chicago, IL: Open Court Publishing.

Gächter, S., D. Nosenzo, and M. Sefton (2008, August). The impact of social comparisons on reciprocity. Working paper, IZA.

Gächter, S. and A. Falk (2002). Reputation and Reciprocity: Consequences for the Labour Relation. *Scandinavian Journal of Economics 104*(1), 1–27.

Gächter, S. and E. Fehr (1999). Collective action as a social exchange. *Journal of Economic Behavior and Organization 39*(4), 341–369.

Gintis, H. (2000). Strong reciprocity and human sociality. *Journal of Theoretical Biology 206*(2), 169–179.

Gintis, H., S. Bowles, R. Boyd, and E. Fehr (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior 24*(3), 153–172.

Greenberg, J. (1990). Employee theft as a reaction to underpayment inequity: The hidden cost of pay cuts. *Journal of Applied Psychology 75*(5), 561–568.

Greiner, B. (2004). The Online Recruitment System ORSEE: A Guide for the Organization of Experiments in Economics. In K. Kremer and V. Macho (Eds.), *Forschung und wissenschaftliches Rechnen 2003*, Göttingen. Ges. für Wiss. Datenverarbeitung. GWDG Bericht 63.

Grišk̆ova, I., O. Rukšėnas, K. Dapšys, S. Herpertz, and J. Höppner (2007). The effects of 10 Hz repetitive transcranial magnetic stimulation on resting EEG power spectrum in healthy subjects. *Neuroscience Letters 419*(2), 162–167.

Gul, F. and W. Pesendorfer (2007). Welfare without happiness. *American Economic Review 97*(2), 471–476.

Gul, F. and W. Pesendorfer (2008). The case for mindless economics. In A. Caplin and A. Schotter (Eds.), *The foundations of Positive and nor-*

mative Economics: A handbook, pp. 3–42. New York City, NY: Oxford University Press, USA.

Hagen, E. and P. Hammerstein (2006). Game theory and human evolution: A critique of some recent interpretations of experimental games. *Theoretical Population Biology 69*(3), 339–348.

Haley, K. and D. Fessler (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior 26*(3), 245–56.

Hennig-Schmidt, H., B. Rockenbach, and A. Sadrieh (forthcoming). In Search of Workers' Real Effort Reciprocity – A Field and a Laboratory Experiment. *Journal of the European Economic Association*.

Henrich, J. and R. Boyd (2001). Why People Punish Defectors: Weak Conformist Transmission can Stabilize Costly Enforcement of Norms in Cooperative Dilemmas. *Journal of Theoretical Biology 208*(1), 79–89.

Homans, G. (1961). *Social behaviour: Its elementary forms*. London, UK: Taylor & Francis.

Houser, D., J. Wooders, and M. Hall (2006). Reputation in auctions: Theory, and evidence from eBay. *Journal of Economics & Management Strategy 15*, 353–369.

Howitt, P. (2002). Looking inside the labor market: a review article. *Journal of Economic Literature 40*(1), 125–138.

Izuma, K., D. Saito, and N. Sadato (2008). Processing of social and monetary rewards in the human striatum. *Neuron 58*(2), 284–294.

Kahneman, D., J. Knetsch, and R. Thaler (1986). Fairness as a constraint on profit seeking: Entitlements in the market. *American Economic Review 76*(4), 728–741.

Keser, C. and F. Van Winden (2000). Conditional cooperation and voluntary contributions to public goods. *Scandinavian Journal of Economics 102*(1), 23–39.

King-Casas, B., D. Tomlin, C. Anen, C. Camerer, S. Quartz, and P. Montague (2005). Getting to know you: reputation and trust in a two-person economic exchange. *Science 308*(5718), 78–83.

Knoch, D., A. Pascual-Leone, K. Meyer, V. Treyer, and E. Fehr (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science 314*(5800), 829–832.

Knoch, D., F. Schneider, D. Schunk, M. Hohmann, and E. Fehr (2009). Disrupting the prefrontal cortex diminishes the human ability to build a good reputation. *Proceedings of the National Academy of Sciences 106*(49), 20895–20899.

Koch, G., M. Oliveri, S. Torriero, G. Carlesimo, P. Turriziani, and C. Caltagirone (2005). rTMS evidence of different delay and decision processes in a fronto-parietal neuronal network activated during spatial working memory. *Neuroimage 24*(1), 34–39.

Kochan, T. and T. Barocci (1985). *Human resource management and industrial relations: text, readings, and cases*. New York City, NY: Little, Brown.

Kocher, M., T. Cherry, S. Kroll, R. Netzer, and M. Sutter (2008). Conditional cooperation on three continents. *Economics Letters 101*(3), 175–178.

Kole, S. and K. Lehn (2000). Workforce Integration and the Dissipation of Value in Mergers. The Case of USAir's Acquisition of Piedmont Aviation. In S. N. Kaplan (Ed.), *Mergers and productivity*, pp. 239–279. University of Chicago Press.

Kreps, D., P. Milgrom, J. Roberts, and R. Wilson (1982). Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic Theory 27*, 245–252.

Kube, S., M. A. Maréchal, and C. Puppe (2008, July). The Currency of Reciprocity – Gift-Exchange in the Workplace. Working Paper 377, IEW.

Kube, S., M. A. Maréchal, and C. Puppe (2010, January). Do Wage Cuts damage Work Morale? Evidence from a Natural Field Experiment. Working Paper 471, IEW.

Kurzban, R., P. DeScioli, and E. O'Brien (2007). Audience effects on moralistic punishment. *Evolution and Human Behavior 28*(2), 75–84.

Kurzban, R. and D. Houser (2005). Experiments investigating cooperative types in humans: A complement to evolutionary theory and simulations. *Proceedings of the National Academy of Sciences 102*(5), 1803–1807.

Landsburg, S. (2007). The Methodology of Normative Economics. *Journal of Public Economic Theory 9*(5), 757–769.

Lawler, E. (1990). *Strategic pay: Aligning organizational strategies and pay systems*. San Francisco, CA: Jossey-Bass.

Lee, D. and N. Rupp (2007). Retracting a gift: How does employee effort respond to wage reductions? *Journal of Labor Economics 25*(4), 725–761.

Levine, D. (1998). Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics 1*(3), 593–622.

Levitt, S. and J. List (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives 21*(2), 153–174.

Loewenstein, G. and E. Haisley (2008). The economist as therapist: methodological ramifications of 'light' paternalism. In A. Caplin and A. Schotter

(Eds.), *The Foundations of Positive and Normative Economics: A Handbook*, pp. 210–245. New York City, NY: Oxford University Press, USA.

Mas, A. and E. Moretti (2009). Peers at work. *American Economic Review 99*(1), 112–145.

Milinski, M., D. Semmann, T. Bakker, and H. Krambeck (2001). Cooperation through indirect reciprocity: image scoring or standing strategy? *Proceedings of the Royal Society B: Biological Sciences 268*(1484), 2495–2501.

Milinski, M., D. Semmann, and H. Krambeck (2002). Reputation helps solve the 'tragedy of the commons'. *Nature 415*(6870), 424–426.

Miller, E. K. and J. D. Cohen (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience 24*, 167–202.

Nowak, M. A. and K. Sigmund (1998). Evolution of indirect reciprocity by image scoring. *Nature 393*, 573–577.

Nowak, M. A. and K. Sigmund (2005). Evolution of indirect reciprocity. *Nature 437*(7063), 1291–1298.

Panchanathan, K. and R. Boyd (2004). Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature 432*(7016), 499–502.

Perugini, M., M. Gallucci, F. Presaghi, and A. Ercolani (2003). The personal norm of reciprocity. *European Journal of Personality 17*(4), 251–283.

Plassmann, H., J. O'Doherty, and A. Rangel (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *Journal of Neuroscience 27*(37), 9984–9988.

Pollis, N. (1968). Reference group re-examined. *The British Journal of Sociology 19*(3), 300–307.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review 83*(5), 1281–1302.

Rege, M. and K. Telle (2004). The Impact of Social Approval and Framing. *Journal of Public Economics 88*(7–8), 1625–1644.

Rigdon, M., K. Ishii, M. Watabe, and S. Kitayama (2009). Minimal social cues in the dictator game. *Journal of Economic Psychology 30*, 358–367.

Robertson, E. M., H. Théoret, and A. Pascual-Leone (2003). Studies in cognition: The problems solved and created by transcranial magnetic stimulation. *Journal of Cognitive Neuroscience 15*(7), 948–960.

Rockenbach, B. and M. Milinski (2006). The efficient interaction of indirect reciprocity and costly punishment. *Nature 444*(7120), 718–723.

Runciman, W. (1966). *Relative deprivation and social justice: A study of attitudes to social inequality in twentieth-century England*. Berkeley, CA: University of California Press.

Sack, A. and D. Linden (2003). Combining transcranial magnetic stimulation and functional imaging in cognitive brain research: possibilities and limitations. *Brain Research Reviews 43*(1), 41–56.

Salpukas, A. (1987, July 21). The 2-tier wage system is found to be 2-edged sword by industry. *The New York Times*, pp. 1 and 47.

Seinen, I. and A. Schram (2006). Social status and group norms: Indirect reciprocity in a repeated helping experiment. *European Economic Review 50*(3), 581–602.

Shapiro, C. and J. Stiglitz (1984). Equilibrium unemployment as a worker discipline device. *American Economic Review 74*(3), 433–444.

Shearer, B. (2003). Compensation Policy and Worker Performance: Identifying Incentive Effects from Field Experiments. *Journal of the European Economic Association 1*(2-3), 503–511.

Shearer, B. (2004). Piece rates, fixed wages and incentives: Evidence from a field experiment. *Review of Economic Studies 71*(2), 513–534.

Simpson, B. and R. Willer (2008). Altruism and indirect reciprocity: The interaction of person and situation in prosocial behavior. *Social Psychology Quarterly 71*(1), 37–52.

Smith, A. (2010). *The Theory of Moral Sentiments*. London, UK: Penguin Classics.

Smith, F., L. Debruine, B. Jones, D. Krupp, L. Welling, and C. Conway (2009). Attractiveness qualifies the effect of observation on trusting behavior in an economic game. *Evolution and Human Behavior 30*(6), 393–397.

Spitzer, M., U. Fischbacher, B. Herrnberger, G. Grön, and E. Fehr (2007). The neural signature of social norm compliance. *Neuron 56*(1), 185–196.

Strobel, A., A. Beauducel, S. Debener, and B. Brocke (2001). Eine deutschsprachige Version des BIS/BAS-Fragebogens von Carver und White. *Zeitschrift für Differentielle und Diagnostische Psychologie 22*(3), 216–227.

Takahashi, H., M. Kato, M. Matsuura, M. Koeda, N. Yahata, T. Suhara, and Y. Okubo (2008). Neural correlates of human virtue judgment. *Cerebral Cortex 18*(8), 1886–1891.

Thaler, R. and C. Sunstein (2003). Libertarian paternalism. *American Economic Review 93*(2), 175–179.

Thöni, C. and S. Gächter (2009, November). Social Comparison and Performance: Experimental Evidence on the Fair Wage-Effort Hypothesis. Working Paper 29, University of St. Gallen, Department of Economics.

Torgler, B., M. Schaffner, B. Frey, S. Schmidt, and U. Dulleck (2008, October). Inequality Aversion and Performance in and on the Field. Working paper, National Centre for Econometric Research.

Walsh, V. and A. Cowey (2000). Transcranial magnetic stimulation and cognitive neuroscience. *Nature Reviews Neuroscience 1*(1), 73–80.

Wassermann, E. (1998). Risk and safety of repetitive transcranial magnetic stimulation: report and suggested guidelines from the International Workshop on the Safety of Repetitive Transcranial Magnetic Stimulation, June 5–7, 1996. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section 108*(1), 1–16.

Wedekind, C. and M. Milinski (2000). Cooperation through image scoring in humans. *Science 288*(5467), 850–852.

Wout, M., R. Kahn, A. Sanfey, and A. Aleman (2005). Repetitive transcranial magnetic stimulation over the right dorsolateral prefrontal cortex affects strategic decision-making. *Neuroreport 16*(16), 1849–1852.

Curriculum Vitae

Frédéric-Guillaume Schneider

Date of birth	October 12, 1981
Place of birth	Werne, Germany
Nationality	German, French

Education

Jan 2006 – present	University of Zürich, Switzerland Doctoral studies in Economics
Jan 2008 – Jan 2009	Foundation of the Swiss National Bank Gerzensee, Switzerland Swiss Program for Beginning Doctoral Students in Economics
Apr 2002 – Aug 2005	University of Bonn, Germany Diplom-Volkswirt (graduate degree in economics)

Professional Experience

Jan 2006 – present	Institute for Empirical Research in Economics, University of Zürich research associate
Oct 2005 – Dec 2005	Centre for European Economic Research, Mannheim, Germany research internship
Oct 2003 – Mar 2004	University of Bonn, Germany teaching assistant in macroeconomics